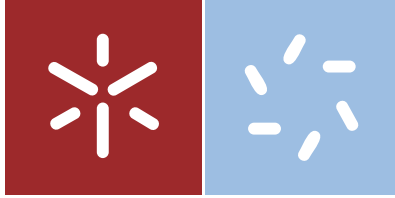


Universidade do Minho  
Escola de Ciências

Liliana Filipa Machado da Silva Faria

Alunos Portugueses no PISA:  
Modelos de Regressão Multinível





Universidade do Minho  
Escola de Ciências

Liliana Filipa Machado da Silva Faria

Alunos Portugueses no PISA:  
Modelos de Regressão Multinível

Tese de Mestrado  
Mestrado em Estatística

Trabalho efectuado sob a orientação de  
Professora Doutora Susana Faria  
Professora Doutora Conceição Portela

## Anexo 3

### DECLARAÇÃO

Nome Liliana Filipa Machado da Silva Faria

Endereço electrónico: scp\_faria\_@hotmail.com

Telefone: 932887533

Número do Bilhete de Identidade: 13018162

Título dissertação /tese

**"Alunos Portugueses no PISA: Modelos de Regressão Multinível"**

Orientadores:

Professora Doutora Susana Faria

Professora Doutora Conceição Portela

Ano de conclusão: 2013/2014

Designação do Mestrado: Mestrado em Estatística

Nos exemplares das teses de doutoramento ou de mestrado ou de outros trabalhos entregues para prestação de provas públicas nas universidades ou outros estabelecimentos de ensino, e dos quais é obrigatoriamente enviado um exemplar para depósito legal na Biblioteca Nacional e, pelo menos outro para a biblioteca da universidade respectiva, deve constar uma das seguintes declarações:

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, 31/10/2014

Assinatura: \_\_\_\_\_

## **Agradecimentos**

Este árduo e longo trabalho deve muito a muitas pessoas, por diferentes razões e, em particular, gostaria de agradecer:

Às minhas orientadoras, Professora Doutora Susana Faria e Professora Doutora Conceição Portela pela orientação, amizade, disponibilidade, palavras de apoio e interesse demonstrados que foram cruciais na realização desta dissertação.

Aos meus pais pelo apoio constante. Agradeço a forma como me educaram e os valores que me transmitiram, a eles dedicarei tudo de bom que a vida me conceder.

Ao meu marido, pela compreensão, força e motivação que sempre me transmitiu e pela muita paciência que teve comigo. Claro que não podia faltar um agradecimento muito especial ao meu pequenote pela companhia que me fez nas longas tardes de trabalho da mamã.

Aos meus colegas e amigos pela amizade, alegria, partilha, companheirismo revelado ao longo destes anos.



## Resumo

Nas últimas décadas houve uma massificação do ensino, o que incrementou a heterogeneidade e a diversidade cultural dos alunos com interesses e motivações diversas, provocando a formação de turmas com características bastantes díspares entre elas e entre os seus elementos constituintes, onde as motivações, os interesses, os níveis sócio culturais, as aptidões e os recursos são diferentes e muito diversificados.

Em muitos trabalhos científicos, as populações estudadas têm uma estrutura hierárquica, ou seja, os indivíduos em estudo estão agrupados em unidades de nível mais baixo, que por sua vez pertencem a unidades de um nível mais alto e assim sucessivamente.

Quanto à educação, tem-se um exemplo clássico: os alunos descritos por notas, estatuto social, idade, entre outros pertencem às turmas (e cada aluno não pertence a mais do que uma turma) que, por sua vez podem ser caracterizadas pelas médias das notas, pelos professores, entre outros, e as turmas pertencem às escolas (e cada turma não faz parte de mais do que uma escola). As escolas estão agrupadas nas Direções Regionais de Educação que por sua vez dependem, em última análise, do Ministério da Educação.

Os modelos multinível são frequentemente utilizados na análise de dados com estrutura hierárquica.

Neste trabalho pretende-se aplicar os modelos de regressão logística multinível a dados de alunos (nível 1) e de escolas nacionais (nível 2), com o objetivo de identificar os principais fatores que podem determinar os resultados escolares dos alunos portugueses relativamente à literacia em matemática e ciências.

Estes modelos foram aplicados aos dados dos alunos portugueses obtidos no âmbito dos testes administrados pelo Programa Internacional de Avaliação de Estudantes (PISA) 2009 na disciplina de matemática e ciências da natureza.

**Palavras-chave:** Desempenho escolar dos alunos; PISA; Modelos multinível; Modelo logístico multinível;





## **Abstract**

In the last few decades there has been a mass education, which increased the heterogeneity and cultural diversity of students with diverse interests and motivations, causing the formation of classes with quite disparate characteristics among them and among their constituents, where the motivations, the interests, the socio-cultural levels, the skills and resources are different and very diverse.

In many scientific studies, the study populations have a hierarchical structure, therefore, the subjects in study are grouped in units of lower level, which in turn belong to a higher level and so on.

Regarding education, we have a classic example: the students are described, among others, by their grades, social status and age, belong to the classes (each student does not belong to more than one class) which in turn can be characterized, among others, by their students average grades and by teachers. The classes belong to schools (and each class is not part of more than one school) they are grouped into Regional Directorates of Education which in turn depend, ultimately, on the Ministry of Education.

Multilevel models are often used in the analysis of data with hierarchical structure.

In this work we intend to apply the multilevel logistic regression models to data from students (level 1) and national schools (level 2), the aim is to identifying the key factors that may determine the educational achievement of Portuguese pupils in relation to mathematics and science literacy.

These models were applied to the Portuguese students data obtained from the tests administered by the Programme for International Student Assessment (PISA) 2009 in the school subjects of mathematics and natural sciences.

**Key words:** Students school performance; PISA; Multilevel models; Multilevel logistic model;



## Conteúdo

Conteúdo.....	ix
Siglas.....	xi
Lista de Tabelas.....	xiii
Lista de Figuras.....	xvi
Capítulo I – Introdução.....	<b>1</b>
Capítulo II – Desempenho Escolar.....	<b>3</b>
2  Fatores que influenciam o desempenho escolar dos alunos.....	3
2.1  Fatores que influenciam o desempenho escolar dos alunos.....	3
2.2  Fatores associados à escola.....	5
Capítulo III – Programa Internacional de Avaliação de Estudantes ( <i>PISA</i> ).....	<b>7</b>
3.1  Objetivos e Países participantes.....	7
3.2  Domínios em Avaliação no PISA: leitura, matemática e ciências.....	9
3.3  População e Amostra.....	11
3.4  Instrumentos de recolha: questionários, exercícios para a avaliação direta de competências.....	13
3.5  Literacia científica.....	15
3.5.1  Interpretação dos níveis de proficiência.....	16
3.6  Literacia matemática.....	16
3.6.1  Interpretação dos níveis de proficiência.....	17
Capítulo IV – Modelo Multinível.....	<b>23</b>
4.1  Modelo de regressão linear.....	23
4.2  Modelos multinível.....	24
4.2.1  Modelos de regressão linear multinível com dois níveis.....	25
4.3  Modelo de regressão logística multinível.....	28
4.3.1  Modelo Nulo.....	29
4.3.2  Modelo de efeitos fixos.....	30
4.4  Métodos de Estimação.....	31
4.4.1  Método – A quasi-verosimilhança (QVP).....	31

4.4.2   Método – A integração numérica ou quadratura de Gauss-Hermite (QGH) .....	32
4.5   Testes Estatísticos .....	32
4.5.1   Teste de Wald .....	32
4.6   Qualidade de ajustamento.....	33
Capítulo V – Análise Exploratória dos dados .....	<b>35</b>
5.1   Dados e Variáveis .....	35
5.2   Análise Exploratória.....	42
5.2.1   Relação entre as variáveis resposta e as variáveis explicativas .....	53
Capítulo VI – Apresentação dos resultados .....	<b>57</b>
6.1   Apresentação dos resultados .....	57
6.2   Apresentação de resultados para a variável resposta (CN) .....	57
6.2.1   Modelo Nulo .....	57
6.2.2   Modelo de regressão logística multinível de efeitos fixos – variáveis explicativas ao nível do aluno .....	58
6.2.3   Modelo de regressão logística multinível de efeitos fixos – variáveis explicativas ao nível do aluno e ao nível da escola .....	61
6.3   Apresentação de resultados para a variável resposta (MAT).....	64
6.3.1   Modelo Nulo .....	64
6.3.2   Modelo de regressão logística multinível de efeitos fixos – variáveis explicativas ao nível do aluno .....	65
6.3.3   Modelo de regressão logística multinível de efeitos fixos – variáveis explicativas ao nível do aluno e ao nível da escola .....	69
Capítulo VII – Conclusão.....	<b>73</b>
Bibliografia.....	<b>77</b>
Anexo.....	<b>83</b>
A1   CN - Regressão logística simples para a variável resposta .....	83
A2   MAT - Regressão logística simples para a variável resposta .....	86

## Siglas

AEC – Índice de atividades extracurriculares;  
AGE – Idade do aluno;  
AGQ – Quadratura Gaussiana Adaptativa;  
AGREG – Constituição do agregado familiar;  
AMU - Área Medianamente Urbana;  
APR - Área Predominantemente Rural;  
APU - Área Predominantemente Urbana;  
CICLOS – Níveis de ensino da escola;  
COMPA – Comportamento do aluno;  
COMPP – Comportamento do professor;  
CN – Nível de desempenho a ciências;  
CTN – Existência de turmas de nível;  
DIM – Dimensão da escola;  
ESCS – Índice económico, social e cultural;  
ESCSMED – Média do índice económico, social e cultural;  
ESTC – Índice do uso de estratégias de controlo;  
ESTCR – Índice do uso de estratégias de compreender e relembrar;  
ESTE – Índice do uso de estratégias de elaboração;  
ESTM – Índice do uso de estratégias de memorização;  
ESTS – Índice do uso de estratégias de estruturação;  
EXPAL – Expectativas do aluno;  
FPCN – Falta de professores de ciências;  
FPMAT – Falta de professores de matemática;  
GAVE – Gabinete de Avaliação Educacional, Ministério da Educação;  
GQ – Quadratura Gaussiana ordinária;  
HEXPCN – Nº médio de explicação a ciências;  
HEXPMAT – Nº médio de explicação a matemática;  
ICC – Coeficiente de correlação intra-classe;

IMI – Situação de imigração;

LOC – Localização onde a escola se encontra;

MAT – Nível de desempenho a matemática;

MQL – Quasi-verosimilhança marginal;

NEF – Nível de ensino que frequenta;

OCDE – Organização de Cooperação e de Desenvolvimento Económico;

OR – Razão de chance;

PCNET – Proporção de computadores ligados à internet;

PISA – Programa de Avaliação Internacional de Estudantes;

PQL – Verosimilhança preditiva ou penalizada;

PRAP – Percentagem de raparigas;

QGH – Método – a integração numérica ou quadratura de Gauss-Hermite;

QVP – Método – a quasi-verosimilhança;

RAP – Rácio aluno/professor;

RPCA – Rácio computador/aluno;

REG – Região a que cada aluno pertence;

REP – Indica se o aluno já repetiu, pelo menos uma vez, um nível de ensino;

SEL – Admissão dos alunos na escola;

SEX – Sexo do aluno;

TIPO – Tipo de escola que o aluno frequenta;

TPC – Ajuda nos trabalhos de casa.

## Lista de Tabelas

Tabela 1   Ciclos do PISA (2000-2009): número de países participantes, alunos e escolas.....	9
Tabela 2   Literacia em leitura .....	9
Tabela 3   Literacia matemática .....	9
Tabela 4   Literacia em científica .....	10
Tabela 5   Amostra de alunos em Portugal segundo as variáveis de estratificação das escolas (2009) .....	10
Tabela 6   Distribuição da amostra em Portugal, por sexo e ano de escolaridade (2009) .....	12
Tabela 7   Evolução das dimensões da literacia científica .....	13
Tabela 8   Descrições dos níveis de proficiência em literacia científica (2009).....	18
Tabela 9   Evolução das dimensões da literacia em matemática.....	19
Tabela 10   Descrições dos níveis de proficiência da literacia matemática (2009).....	20
Tabela 11   Estatísticas descritivas das variáveis explicativas contínuas .....	21
Tabela 12   Distribuição dos alunos.....	44
Tabela 13   Distribuição das escolas.....	45
Tabela 14   Tabela de contingência entre as variáveis resposta CN e MAT e as variáveis explicativas qualitativas ao nível do aluno.....	46
Tabela 15   Teste de independência entre as variáveis resposta CN e MAT e as variáveis explicativas qualitativas ao nível do aluno.....	48
Tabela 16   Estatísticas descritivas das variáveis explicativas contínuas por nível da variável resposta CN .....	49
Tabela 17   Estatísticas descritivas das variáveis explicativas contínuas por nível da variável resposta MAT .....	50
Tabela 18   Teste F e Teste t entre as variáveis resposta CN e MAT e as variáveis explicativas contínuas ao nível do aluno .....	50
Tabela 19   CN - Estimativa da variância (modelo nulo).....	51
Tabela 20   CN - Estimativa do efeito fixo .....	57
Tabela 21   CN - Modelo de efeitos fixos com variáveis explicativas ao nível do aluno .....	58

Tabela 22  CN - Modelo de efeitos fixos com variáveis explicativas ao nível do aluno e da escola.....	59
Tabela 23  MAT - Estimativa da variância (modelo nulo).....	62
Tabela 24  MAT - Estimativa do efeito fixo .....	64
Tabela 25  MAT - Modelo de efeitos fixos com variáveis explicativas ao nível do aluno .....	66
Tabela 26  MAT - Modelo de efeitos fixos com variáveis explicativas ao nível do aluno e da escola.....	70
A1  CN - Regressão logística simples com a variável explicativa <i>SEX</i> .....	83
A1  CN - Regressão logística simples com a variável explicativa <i>AGE</i> .....	83
A1  CN - Regressão logística simples com a variável explicativa <i>IMI</i> .....	83
A1  CN - Regressão logística simples com a variável explicativa <i>NEF</i> .....	83
A1  CN - Regressão logística simples com a variável explicativa <i>AGREG</i> .....	83
A1  CN - Regressão logística simples com a variável explicativa <i>TPC</i> .....	84
A1  CN - Regressão logística simples com a variável explicativa <i>HEXPCN</i> .....	84
A1  CN - Regressão logística simples com a variável explicativa <i>EXPAL</i> .....	84
A1  CN - Regressão logística simples com a variável explicativa <i>ESTC</i> .....	84
A1  CN - Regressão logística simples com a variável explicativa <i>ESTS</i> .....	84
A1  CN - Regressão logística simples com a variável explicativa <i>ESTE</i> .....	85
A1  CN - Regressão logística simples com a variável explicativa <i>ESTM</i> .....	85
A1  CN - Regressão logística simples com a variável explicativa <i>ESTCR</i> .....	85
A1  CN - Regressão logística simples com a variável explicativa <i>REG</i> .....	85
A1  CN - Regressão logística simples com a variável explicativa <i>REP</i> .....	85
A1  CN - Regressão logística simples com a variável explicativa <i>ESCS</i> .....	85
A2  MAT - Regressão logística simples com a variável explicativa <i>SEX</i> .....	86
A2  MAT - Regressão logística simples com a variável explicativa <i>AGE</i> .....	86
A2  MAT - Regressão logística simples com a variável explicativa <i>IMI</i> .....	86
A2  MAT - Regressão logística simples com a variável explicativa <i>NEF</i> .....	86
A2  MAT - Regressão logística simples com a variável explicativa <i>AGREG</i> .....	86
A2  MAT - Regressão logística simples com a variável explicativa <i>TPC</i> .....	87
A2  MAT - Regressão logística simples com a variável explicativa <i>HEXPCN</i> .....	87



A2   MAT - Regressão logística simples com a variável explicativa <i>EXPAL</i> .....	87
A2   MAT - Regressão logística simples com a variável explicativa <i>ESTC</i> .....	87
A2   MAT - Regressão logística simples com a variável explicativa <i>ESTS</i> .....	87
A2   MAT - Regressão logística simples com a variável explicativa <i>ESTE</i> .....	88
A2   MAT - Regressão logística simples com a variável explicativa <i>ESTM</i> .....	88
A2   MAT - Regressão logística simples com a variável explicativa <i>ESTCR</i> .....	88
A2   MAT - Regressão logística simples com a variável explicativa <i>REG</i> .....	88
A2   MAT - Regressão logística simples com a variável explicativa <i>REP</i> .....	88
A2   MAT - Regressão logística simples com a variável explicativa <i>ESCS</i> .....	88

## Lista de Figuras

Figura 1   Percentagem de alunos nos dois níveis de desempenho as disciplinas de ciências e matemática.....	43
Figura 2   Ajuda nos trabalhos de casa vs género do aluno .....	52
Tabela 3   Tipo de escola vs divisão regional .....	52
Figura 4   Níveis de desempenho a ciências vs horas de explicação a ciências .....	53
Figura 5   Níveis de desempenho a matemática vs horas de explicação a matemática .....	54
Tabela 6   Níveis de desempenho a ciências vs tipo de escola .....	54
Figura 7   Figura 7   Níveis de desempenho a matemática vs tipo de escola.....	55
Tabela 8   Níveis de desempenho a ciências e matemática vs género do aluno .....	58

# Capítulo I

## Introdução

*“A tomada de consciência da sociedade em geral para a importância da formação escolar e o desenvolvimento económico e industrial do país nas últimas décadas vieram originar, por consequência, o alargamento da escolaridade obrigatória em Portugal, em 1986. Esta medida contribuiu para que a heterogeneidade e a diversidade cultural fossem uma realidade na população escolar”* (Cortesão, 1998).

Na escola já não se encontram só os filhos das famílias economicamente mais favorecidas, ou os alunos que estão motivados para a obtenção de uma escolaridade superior, encontram-se também, todo um conjunto de alunos, filhos de famílias mais carenciadas, com outros interesses e motivações que, portanto, dão origem a turmas bastante heterogéneas onde os interesses, os níveis sócio económicos e culturais, as aptidões e os recursos são diferentes e diversificados.

No entanto a igualdade de acesso escolar não é, geralmente, acompanhada por uma igualdade de sucesso havendo vários fatores que contribuem e influenciam o desempenho escolar (Cortesão, 1998).

O Programa de Avaliação Internacional de Estudantes (PISA) da Organização de Cooperação e de Desenvolvimento Económico (OCDE) avalia o desempenho dos alunos na literacia em leitura, matemática e ciências, no que se refere à compreensão de conceitos fundamentais, ao domínio de certos processos e à capacidade dos jovens usarem os seus conhecimentos e as suas competências na resolução de problemas da vida real. Os resultados obtidos com este programa tem uma grande utilidade na medida em que fornecem dados comparáveis ao longo do tempo e entre países, os quais permitem a avaliação dos sistemas educativos e, implicitamente, do retorno obtido com a despesa em educação.

O universo do *PISA* é constituído pela população de estudantes com cerca de 15 anos que frequentam as escolas num determinado país onde são residentes. Em Portugal, participaram no programa, em 2009, 6298 alunos pertencentes a 214 escolas.

Na Educação, as populações em estudo têm uma estrutura hierárquica, por níveis, isto é: alunos, turmas, escolas, (entre outros), que constituem uma sequência natural de

agrupamentos hierárquicos, de tal forma que as variáveis de um nível podem interagir com outras variáveis dentro do mesmo nível ou de outro nível.

Os modelos multinível são uma resposta à necessidade de analisar a relação entre os indivíduos e o meio que os rodeia. Através destes modelos podemos separar o papel de cada uma das características de uma estrutura interativa complexa, com o intuito de melhorar o conhecimento da realidade, permitindo uma intervenção mais eficiente.

O objetivo principal do trabalho é identificar os principais fatores que podem influenciar o desempenho escolar dos alunos portugueses relativamente à literacia em matemática e em ciências usando os dados obtidos no âmbito do Programa Internacional de avaliação de Estudantes (PISA) 2009.

Esta tese divide-se em sete capítulos: o Capítulo I – Introdução, onde se faz uma breve apresentação do tema a abordar na dissertação; Capítulo II – Desempenho Escolar, onde são apresentados os vários fatores que influenciam o desempenho escolar dos alunos; Capítulo III – Programa Internacional de Avaliação de Estudantes (PISA), onde se descreve o programa de PISA; Capítulo IV – Modelo Multinível, onde se desenvolvem os conteúdos teóricos e respetivas metodologias; Capítulo V – Análise Exploratória, onde a análise descritiva das variáveis em estudo é realizada, Capítulo VI – Apresentação dos Resultados, onde a aplicação da metodologia aos dados obtidos no âmbito do PISA 2009 é apresentada; por fim o Capítulo VII – Conclusão e Trabalho Futuro, onde se apresentam as conclusões finais do trabalho desenvolvido bem como linhas de investigação para trabalhos futuros.

## Capítulo II

### Desempenho Escolar

Citando Garcia (1995): *“ Ensinar é algo que qualquer um faz em qualquer momento, mas isso não é o mesmo que ser professor (...) o professor eficaz é um ser humano único que aprendeu a fazer uso de si próprio eficazmente, e a realizar os seus propósitos e os da sociedade na educação de outras pessoas”*

## 2| Fatores que influenciam o desempenho escolar dos alunos

Muitos estudos na área da educação mostram que são vários os fatores que podem influenciar o desempenho escolar dos alunos.

Estes fatores podem ser classificadas em três grandes grupos: os associados à família (características socioeconómicas e culturais), aqueles relacionados ao próprio aluno (habilidade, motivação) e os associados a fatores escolares (infra-estruturas, práticas didáticas, características dos professores) (Ferrão e Fernandes, 2000), (Soares, 2004).

### 2.1| Fatores associados à família e ao próprio aluno

Diversos sociólogos realizaram estudos nesta área. Durante os anos de 1950 e 1960, em países como os Estados Unidos, Inglaterra e França, concluiu-se em vários trabalhos de investigação que os resultados escolares encontram-se diretamente relacionados com as características socioeconómicas e culturais dos alunos (Coleman et al., 1966).

A família, o seu interesse e o seu estímulo, assim como a quantidade de incentivo e motivação e de interação linguística, têm um papel preponderante na adaptação social da criança e nos seus resultados escolares (Fonseca, 2004), (Jiménez, 1997).

Geralmente, os pais têm um conjunto de expectativas em relação ao futuro dos seus filhos. Estas expectativas não são as mesmas em todas as classes sociais. Além de diferirem nos projetos que imaginam para os filhos, são também encontradas diferenças na imagem que os pais têm acerca da escola. Esta vai ser de grande relevância na forma como a própria

criança se situará face à escola, à aprendizagem e à importância da mesma para si e, por consequência, sobre o seu futuro nesta instituição (Pierre et al., 1994).

A família molda a criança segundo as suas experiências educativas e os seus hábitos culturais. Por outro lado, a família determina a evolução da criança e a boa qualidade deste desenvolvimento, comprovada pela escola, traduz-se em bons resultados escolares (Martins e Cabrita, 1991).

A família é considerada o foco central, isto é, é o espaço educativo para o desenvolvimento moral, cognitivo e afetivo onde se educam e criam as crianças para que estas construam o seu próprio “eu”. É também na família, espaço histórico e simbólico, que as crianças encontram o espaço favorecido da construção social da realidade em que, através das relações entre os seus membros, as tarefas do quotidiano recebem o seu significado (Diogo, 1998).

*“No geral, as famílias favorecidas tem maiores possibilidades em garantir à criança um desenvolvimento harmonioso e uma inserção fácil nos meios exteriores como a escola. A imagem que os pais têm de si próprio, e do filho, influencia fortemente o seu desenvolvimento. Certos comportamentos que a mãe tem com o filho, como a sensibilidade, a aceitação, a cooperação, a capacidade para exprimir as suas emoções, são favoráveis ou desfavoráveis. Quando esta interação é positiva a criança é mais aberta socialmente, mais independente, capaz de uma atenção suportada, mais segura nas suas experiências. Pelo contrário, quando é negativa, a criança é mais ansiosa, insegura e insatisfeita. As características psicológicas dos pais também assinalam significativamente a adaptação das crianças à sociedade, nomeadamente no ambiente escolar. A qualidade das trocas afetivas depende do equilíbrio emocional vivido na família através das experiências vividas pela criança”* (Pierre et al., 1994).

Os estudos nacionais e internacionais têm concluído que os fatores associados ao contexto socioeconómico familiar dos alunos e às suas capacidades inatas são os mais determinantes na explicação dos seus resultados escolares. Segundo o estudo de Pereira (2011) sobre uma análise da evolução do desempenho dos estudantes portugueses no PISA que se debruça sobre a evolução das classificações dos estudantes portugueses nos ciclos entre 2003 e 2009. Uma análise mais profunda da evolução do desempenho requer, contudo, um

confronto com a evolução das características da população estudantil e das escolas. Desde logo, o estatuto socioeconómico dos estudantes tem influência sobre o desempenho, pelo que a alteração deste estatuto ao longo das sucessivas realizações do programa deverá ser tida em conta.

## **2.2 | Fatores associados à Escola**

Ferrão e Fernandes (2003) definem escola eficaz como uma escola que assume uma parte das responsabilidades nos resultados escolares obtidos pelos alunos. Também Murillo (2003) define que uma escola é eficaz quando consegue um bom desenvolvimento por parte de todos os seus alunos tendo em conta o seu rendimento prévio e a situação social, económica e cultural das famílias.

Os resultados de vários estudos mostram evidências de que o efeito escola pode agir tanto no nível do desempenho escolar do aluno quanto no grau de equidade dentro dos sistemas de ensino ou mesmo dentro das escolas, em relação às diferenças socioeconómicas e culturais dos alunos.

Dronkers e Robert (2003) mostraram com um estudo realizado com o PISA 2000, que quando se controlam as características dos alunos e dos pais, o desempenho escolar dos alunos que frequentam escolas privadas dependentes do governo são melhores do que em escolas públicas. Num estudo posterior Corten e Dronkers (2004), verificaram que o efeito associado às escolas privadas dependentes do governo diferia entre alunos com níveis socioeconómicos diferentes.

Um estudo realizado por Soares, Sátyro e Mambrini (2000) sobre fatores da eficácia escolar conclui que uma escola é eficaz quando possui uma equipa de professores habilitados. Estes autores consideram ainda que a formação do professor está diretamente relacionada com a formação dos alunos.

Gutierrez (2005), aplicando modelos de regressão multinível aos dados do PISA 2000 de alunos do Perú, realizou um estudo para verificar a contribuição da escola no desempenho escolar dos alunos. Concluiu que existe um efeito escola significativo, mesmo quando a proficiência dos alunos é controlada pelo capital cultural e pelo capital económico ao nível

do aluno e pelo capital cultural médio ao nível da escola. Entre os fatores associados às diferenças no desempenho entre as escolas, encontrou como significativos os "recursos educativos", a "infraestrutura física da escola", a "motivação dos professores" e a "falta de professores".

Também Ferrão e Dias (2006) realizaram um estudo com o PISA 2000, onde pretendiam comparar o desempenho dos alunos com um nível socioeconómico desfavorecidos em escolas públicas e privadas. Este estudo revelou que, quando o clima de sala de aula e as condições de ensino aprendizagem são controlados não existem diferenças entre o tipo de escola, ou seja, não existem diferenças no desempenho dos alunos com um nível socioeconómico desfavorecido.

Não obstante, as escolas poderão diferir, por exemplo, quanto à sua organização e ao corpo docente (escolas localizadas em áreas mais desenvolvidas são suscetíveis de atrair melhores professores).



## Capítulo III

### Programa Internacional de Avaliação de Estudantes (PISA)

*“...o principal objetivo do PISA é mostrar em que medida os alunos, no final da escolaridade obrigatória, estarão preparados para fazer face às novas exigências da sociedade”* (GAVE,2010:6)

#### 3.1| Objetivos e países participantes

O Programa Internacional de Avaliação de Estudantes (PISA) é um esforço conjunto empreendido pelos seus participantes – os países membros da OCDE, assim como mais de 30 países parceiros – para avaliar em que medida os alunos aos 15 anos de idade estão preparados para enfrentar os desafios que se lhes deparam na vida futura.

A avaliação do PISA tem como principal objetivo medir o conhecimento, as competências e as atitudes que refletem mudanças em curso nos currículos dos vários países, indo além da abordagem escolar para a utilização dos conhecimentos em tarefas quotidianas. É uma avaliação baseada num modelo dinâmico de aprendizagem ao longo da vida em que novos conhecimentos e capacidades são necessários para uma adaptação bem-sucedida num mundo em constante mudança.

*“O aspeto essencial do PISA é o de assentar numa avaliação incidindo nas competências que evidenciem o que os jovens de 15 anos sabem, valorizam e são capazes de fazer em contextos pessoais, sociais e globais. Esta perspetiva difere das que se baseiam exclusiva e exhaustivamente nos currículos oficiais; no entanto, inclui problemas situados em contextos educativos e profissionais e reconhece o papel essencial do conhecimento, dos métodos, atitudes e valores que definem as disciplinas científicas. A expressão que melhor descreve o objeto de avaliação nas diferentes áreas no PISA é a de literacia”* (GAVE, 2007).

Organizado em ciclos de análise trienais, o PISA recolhe informação sobre conhecimentos específicos e competências dos alunos em três domínios fundamentais: leitura, matemática e ciências. Em cada um dos ciclos de avaliação é eleito um domínio principal. A primeira recolha de informação ocorreu em 2000 e teve como principal domínio de avaliação a

literacia em contexto de leitura. O estudo envolveu, então, cerca de 228 784 alunos de 15 anos, de 32 países, 28 dos quais membros da OCDE. Em Portugal o PISA envolveu 149 escolas (sendo 138 públicas e 11 privadas), abrangendo 4604 alunos, desde o 7º ao 11º ano de escolaridade.

O PISA 2003 contou com 41 países, incluindo a totalidade dos membros da OCDE (30), envolvendo 276 165 alunos de 15 anos. O estudo deu um maior enfoque à literacia matemática. Em Portugal, o PISA envolveu 153 escolas (sendo 141 públicas e 12 privadas), abrangendo 4608 alunos, desde o 7º ao 11º ano de escolaridade.

No estudo PISA que decorreu em 2006, houve preponderância da literacia científica e contou com a participação de cerca de 60 países, envolvendo 398 750 alunos de 7 000 escolas. Em Portugal, o PISA envolveu 173 escolas (sendo 155 públicas e 18 privadas), abrangendo 5109 alunos, desde o 7º ao 11º ano de escolaridade.

No estudo PISA que decorreu em 2009 envolveu 475 460 alunos e contou com a participação de 65 países. Em Portugal, envolveu 212 escolas (sendo 152 públicas e 20 privadas), abrangendo 5109 alunos, desde o 7º ao 11º ano de escolaridade. Os testes PISA foram aplicados a 6298 alunos (214 escolas) e teve como principal domínio de avaliação a literacia em contexto de leitura.

O PISA recolhe ainda dados sobre a escola e o contexto dos alunos em casa, as suas estratégias de aprendizagem, ambientes de aprendizagem e sua familiaridade com computadores.

Até ao momento participam no PISA 66 países, sendo possível distinguir diferentes níveis de regularidade na sua participação.

Em termos globais, a dimensão da amostra de escolas e de alunos tem vindo a aumentar desde 2000, em Portugal e nos restantes países participantes (Tabela 1), com o objetivo de aumentar a precisão e confiança associados aos resultados e possibilitar também a elaboração de análises com base em informação mais desagregada.

Resumidamente, apresenta-se na Tabela 1 o número de alunos participantes no PISA desde 2000 a 2009.

**Tabela 1 | Ciclos do PISA (2000-2009): número de países participantes, alunos e escolas**

	<b>2000</b>	<b>2003</b>	<b>2006</b>	<b>2009</b>
<b>Nº de países participantes (Total)</b>	43	41	57	65
<b>Nº de alunos participantes (Total)</b>	228784	276165	398750	475460
<b>Nº escolas (Portugal)</b>	149	153	173	214
<b>Nº alunos participantes (Portugal)</b>	4585	4608	5109	6298

### 3.2 | Domínios em avaliação no PISA: leitura, matemática e ciências

Como se referiu, o PISA incide em três grandes domínios ou áreas de conhecimento e competências, consideradas decisivas para a vida no século XXI: leitura (Tabela 2), matemática (Tabela 3) e ciências (Tabela 4).

**Tabela 2 | Literacia em leitura**

*Literacia em leitura*

A literacia em leitura remete “a capacidade do indivíduo compreender, usar, refletir sobre e apropriar-se de textos escritos, de forma a alcançar os seus objetivos, desenvolver o próprio conhecimento e potencial e participar na sociedade (OCDE, 2009c)” (GAVE, 2010:6)

**Domínio principal: 2000 e 2009**

**Tabela 3 | Literacia matemática**

*Literacia matemática*

A literacia matemática tem a ver com “a capacidade de um indivíduo identificar e compreender o papel que a matemática desempenha no mundo real, de fazer julgamentos bem fundamentados e de usar e se envolver na resolução matemática de problemas da sua vida, enquanto cidadão construtivo, preocupado e reflexivo (OCDE, 2003a)” (GAVE, 2010:6)

**Domínio principal: 2003 e 2012**

**Tabela 4 | Literacia científica**

<i>Literacia científica</i>
<p>A literacia científica “refere-se ao conhecimento científico, e à utilização desse conhecimento para identificar questões relacionadas com a ciência; à compreensão das características próprias da ciência enquanto forma de conhecimento e de investigação; à consciência do modo a ciência e tecnologia influenciam os ambientes material, intelectual e cultural das sociedades; e à vontade de envolvimento em questões relacionadas com a ciência e com o conhecimento científico, enquanto cidadão consciente (OCDE, 2006)” (GAVE, 2010:6)</p> <p><b>Domínio principal: 2006</b></p>

As três definições apresentadas clarificam que no PISA o enfoque analítico não se centra na mera avaliação de conteúdos curriculares dos estudantes, mas sim numa avaliação dos conhecimentos e competências que lhes permitam agir, de forma consciente e informada, nas sociedades atuais, enfrentando adequadamente os desafios que esta lhes coloca, e tornando-se, assim, cidadãos conscientes, ativos e refletivos.

Assim, o enfoque da avaliação da literacia em leitura no PISA remete, sobretudo, para o uso de textos escritos no dia-a-dia dos alunos, e para conhecimentos sobre a língua, características gramaticais, figuras de estilo, ou compreensão do vocabulário ou das referências culturais incorporadas nos textos. Neste sentido, como se referiu, “o principal objetivo do PISA é mostrar em que medida os alunos, no final da escolaridade obrigatória, estarão preparados para fazer face às novas exigências da sociedade” (GAVE, 2010:6).

Já a avaliação da literacia matemática “ (...) *prende-se principalmente com o uso abrangente e funcional da matemática, com a capacidade de reconhecer e de formular problemas matemáticos em várias situações*” (GAVE, 2010:6).

Por sua vez, a literacia científica remete para a avaliação de competências, de conhecimentos, e de atitudes face a determinadas situações ou contextos (GAVE, 2010:6).

De um modo geral, nos estudos de PISA o conceito de literacia está relacionado com o domínio de processos, compreensão de conceitos e capacidades para atuar em diversas situações (OCDE, 1999).

Para concretizar uma avaliação com estas características, o PISA distingue, analiticamente diferentes tipos de elementos:

- \* **Conteúdos** ou conhecimentos em cada domínio que os alunos precisam de aplicar (forma ou tipo de materiais de leitura na literacia em leitura, conceitos matemáticos relevantes na literacia matemática e conhecimento científico na literacia científica);
- \* **Processos** ou competências que, em cada domínio, os alunos têm de utilizar (tipo de tarefa de leitura na literacia em leitura, clusters de competências na literacia matemática e competências científicas na literacia científica);
- \* **Contextos** para que remetem os vários problemas (uso previsto para o texto na literacia em leitura, situações na literacia matemática e situações que envolvem ciência e tecnologia na literacia científica);
- \* **Atitudes** dos alunos e disposições face à aprendizagem.

### 3.3| População e amostra

A amostra do PISA em Portugal, e na generalidade dos países, é uma mostra estratificada em duas etapas. Na primeira etapa é selecionada uma amostra de escolas, na segunda, uma amostra de alunos nas escolas previamente selecionadas.

Todo o processo de contagem é monitorizado pela OCDE, que assegura diretamente alguns dos procedimentos, garantindo assim que todos os países cumprem os requisitos estabelecidos.

As variáveis para a estratificação da amostra de escolas em Portugal foram acordadas com a OCDE e têm sido idênticas desde 2000 (Tabela 5):

- \* Região – NUTS II;
- \* Tipologia de escola – Escola Básica do 2º e 3º Ciclo / Escola Básica do 3º Ciclo e do Ensino Secundário / Escola Secundária;
- \* Natureza institucional da escola – Público / Privada;
- \* Tipologia das áreas urbanas – Área Medianamente Urbana (AMU), Área Predominantemente Rural (APR) e Área Predominantemente Urbana (APU).

Em Portugal, o número de escolas e alunos participantes em cada estrato garante a representatividade dos resultados do PISA Portugal, por região, por tipologia de escola, por natureza de escola (público e privado) e ainda por tipologia de área urbana.

Na Tabela 6 observa-se que em Portugal, a maioria dos alunos participantes no PISA 2009 encontrava-se a frequentar o 10º ano ou o 9º ano. Mesmo assim, tal como noutros anos, continuam a verificar-se valores relativamente elevados de alunos em graus de escolaridade inferiores ao 9º ano, o que decorrerá certamente de situações de retenção escolar.

**Tabela 5| Amostra de alunos em Portugal segundo as variáveis de estratificação das escolas (2009)**

<b>Região (NUTS II)</b>						
<b>Alentejo</b>	<b>Algarve</b>	<b>Centro</b>	<b>Lisboa</b>	<b>Norte</b>	<b>R.A.Madeira</b>	<b>R.A.Açores</b>
832	790	1400	1161	1922	102	91
<b>Tipologia</b>						
<b>Escolas Básicas</b>		<b>Escolas Básicas e Secundárias</b>		<b>Escolas Secundárias</b>		
1423		4420		455		
<b>Público</b>		<b>Público / Privada</b>		<b>Privado</b>		
5616				682		
<b>Tipologia de área urbana</b>						
<b>Área predominantemente Rural (APR)</b>		<b>Área medianamente Urbana (AMU)</b>		<b>Área predominantemente Urbana (APU)</b>		
95		736		5467		
<b>Total de alunos Participantes = 6298</b>						

Tabela 6 | Distribuição da amostra em Portugal, por sexo e ano de escolaridade (2009)

<i>Ano de Escolaridade</i>				
<i>7º Ano</i>	<i>8º Ano</i>	<i>9º Ano</i>	<i>10º Ano</i>	<i>11º Ano</i>
134	517	1560	3619	25
<i>Sexo</i>				
<i>Feminino</i>		<i>Masculino</i>		
3278		3020		

### 3.4 | Instrumentos de recolha: questionários exercícios para a avaliação direta de competências

São inúmeros os instrumentos de recolha de conhecimento desenvolvidos e aplicados no âmbito do projeto PISA. Assim, para além dos testes para avaliação de competências e conhecimentos em cada um dos domínios (que em seguida se explicitam), fazem parte da operação de inquirição vários inquéritos por questionário, dirigidos a diferentes atores (alunos, escolas, encarregados de educação), que visam complementar a informação obtida. Alguns desses questionários são de resposta obrigatória em todos os países, outros constituem opções, cabendo aos países a decisão de os aplicar, ou não.

São, os seguintes, os inquéritos por questionário aplicados em Portugal, em 2009, no âmbito do PISA:

- \* Questionário base aos alunos. Os alunos respondem a um questionário de cerca de 30 minutos sobre si próprios e o seu contexto familiar. Nele incluem-se quer perguntas de caracterização socioeconómica e cultural dos alunos e suas famílias (contexto familiar, profissional e cultural), quer perguntas sobre a trajetória escolar dos alunos, as suas

atividades de leitura, o tempo dedicado ao estudo, as características da escola, o acesso a bibliotecas e as estratégias de leitura e compreensão de textos;

- \* Questionário aos alunos sobre acesso e uso de tecnologias de informação. Trata-se de um módulo adicional do questionário aos alunos que visa aprofundar informação sobre o uso das tecnologias da informação;
- \* Questionário às escolas. Um questionário de cerca de 20 minutos de duração respondido pelos responsáveis dos Conselhos Executivos das escolas selecionadas;
- \* Questionário aos encarregados de educação. Neste questionário estão incluídas perguntas de caracterização sócio-demográfica, perguntas para avaliar o grau de envolvimento com a leitura, os recursos e hábitos de leitura em casa, as perceções sobre a escola e envolvimento com a mesma, e ainda os critérios de escolha da escola.

Para cumprir o seu principal objetivo, a avaliação das competências e conhecimentos dos alunos em domínios considerados fundamentais, o PISA tem como instrumento de recolha de informação vários exercícios, ou provas, a serem resolvidos pelos alunos. O ponto de partida para os exercícios é sempre um estímulo ou suporte (excertos de textos, gráficos, diagramas, ou uma combinação de vários elementos), a propósito do qual são colocadas várias perguntas. O conjunto constituído pelo estímulo e respetivas perguntas é designado por “unidade” e esta pode ter até cinco perguntas ou itens.

As perguntas colocadas a propósito dos vários suportes ou estímulos podem implicar diferentes modalidades de resposta. Algumas exigem que os alunos selecionem a resposta correta a partir de uma lista de opções que lhes é apresentada; outras têm uma forma de resposta “aberta”, devendo o aluno escrever a sua resposta. Estas últimas são mais exigentes do ponto de vista do apuramento dos resultados, uma vez que implicam que uma equipa de classificadores decida, com base em critérios rigorosos, seguidos por todos os países, se a resposta dada é, ou não, correta.



### 3.5| Literacia científica

Com a reestruturação concetual resultante do ciclo de 2006, a literacia científica passa a ser definida com base em quatro principais dimensões diferenciadas:

- \* **Conteúdos** – conhecimentos dos alunos e a sua capacidade para utilizar efetivamente esses conhecimentos, enquanto executam determinados processos cognitivos característicos da ciência e da investigação científica, em contextos de relevância pessoal, social e global (GAVE: 2007);
- \* **Processos ou competências** – processos científicos, centrados na capacidade de adquirir, interpretar e agir baseado em evidência. O PISA destaca três tipos de competências para as quais é necessária a aplicação tanto de conhecimento do mundo natural, como da própria investigação científica;
- \* **Contextos** – situações da vida quotidiana, e não limitadas ao contexto escolar, que envolvam ciência e tecnologia: pessoal, social e global;
- \* **Atitudes** – a dimensão das atitudes desempenha “um papel significativo no interesse, na atenção e nas reações dos indivíduos face à ciência e à tecnologia em geral” (GAVE, 2007).

Assim, a educação científica assume-se como uma via para o desenvolvimento de atitudes de sensibilização dos alunos para as questões científicas, adquirindo e aplicando conhecimentos científicos e tecnológicos em benefício próprio e da sociedade.

Na Tabela 7 apresenta-se um resumo da evolução no domínio das ciências. Observa-se uma grande alteração ao longo dos ciclos. Do primeiro para o segundo ciclo as principais alterações foram ao nível dos processos. Neste terceiro ciclo, registaram-se alterações a todos os níveis: uma reestruturação dos conteúdos científicos, passando de uma listagem de áreas do conhecimento para uma divisão de dois domínios de conhecimentos principais; uma particularização e simplificação da terminologia das competências científicas; uma maior pormenorização dos contextos de aplicação da ciência face às três grandes categorias de referência nos dois primeiros ciclos; e por fim a introdução das atitudes enquanto nova dimensão da literacia científica. Apesar das alterações referidas, o domínio das ciências acompanha os domínios anteriores na forma como foi cuidadosamente definido pelas

equipas do PISA no sentido de assegurar a comparação metodológica e concetual ao longo dos ciclos.

### 3.5.1 | Interpretação dos níveis de proficiência

O desempenho dos alunos no PISA é apresentado quer sob a forma de uma pontuação quer sob a forma de um nível de proficiência que descreve as tarefas que um dado aluno é capaz de realizar com sucesso.

Assim, no sentido de captar a evolução da complexidade e dificuldade das tarefas e das capacidades dos alunos no PISA 2009, a escala global de ciências foi dividida em seis níveis, cuja descrição detalhada aparece na Tabela 8.

*“É conveniente, numa sociedade movida também pelos rápidos processos de inovação tecnológica, manter um olhar atento às gerações atuais, incluindo os jovens de 15 anos, e identificar os alunos pautados por “excelência”. O estudo do PISA permite fazê-lo ao classificar os alunos por níveis de proficiência, correspondendo os níveis 5 e 6 aos alunos hábeis de um raciocínio científico avançado, o que lhes permite fazer parte de um grupo de futuros inovadores que irão impulsionar as capacidades tecnológicas e de mudanças nas indústrias científicas (OCDE, 2010d: 24).*

### 3.6 | Literacia em matemática

O conceito de literacia matemática é avaliada com base em três principais dimensões:

- \* **Conteúdo matemático:** definido principalmente em termos de quatro categorias (quantidade, espaço e forma, mudanças e relações, indeterminação), e apenas de forma secundária em relação a linhas curriculares, tais como números, álgebra e geometria;
- \* **Processos matemáticos ou competências:** definidos por meio de competências matemáticas gerais, que incluem a utilização de capacidades em linguagem matemática, modelagem e resolução de problemas. As questões são organizadas em conjuntos de competências que definem o tipo de capacidade de pensamento necessário;

- \* **Situações:** definidas em termos da utilização da matemática, com base no seu distanciamento dos estudantes. A estrutura identifica cinco situações: pessoal, educacional, ocupacional, pública e científica.

É possível perceber que este domínio foi o que sofreu menos alteração ao longo dos quatro ciclos do PISA (Tabela 9), na medida em que, tendo ocupado o lugar de domínio primordial em 2003, manteve a sua organização inalterada nos ciclos seguintes.

### **3.6.1 | Interpretação dos níveis de proficiência**

Em 2003, ano em que a matemática assumiu o estatuto de domínio principal, foi possível apresentar, para esta área, descrições bastante detalhadas da proficiência dos alunos. Assim, no sentido de captar a evolução da complexidade e da dificuldade das tarefas e das capacidades dos alunos no PISA, a escala global da matemática foi dividida em seis níveis, cuja descrição detalhada aparece na Tabela 10.

No estudo do PISA, são considerados alunos “de topo” aqueles que registam de nível de 5 ou 6 de proficiência, ou seja, onde a capacidade de lidar com informações complexas é maior (OCDE, 2009c: 121).

Tabela 7 | Evolução das dimensões da literacia científica

	2000	2003	2006	2009
<b>Conteúdos</b> (conhecimento científico)	Treze áreas do conhecimento e conceitos científicos: <ul style="list-style-type: none"> <li>• Biodiversidade</li> <li>• Forças e movimento</li> <li>• Alterações fisiológicas</li> </ul>	Áreas do conhecimento e conceitos científicos: <ul style="list-style-type: none"> <li>• Biodiversidade</li> <li>• Forças e movimento</li> <li>• Alterações fisiológicas</li> </ul>	Sobre o mundo natural: <ul style="list-style-type: none"> <li>• “Sistemas físicos”</li> <li>• “Sistemas vivos”</li> <li>• “Sistemas da terra e do espaço”</li> <li>• “Sistemas de tecnologia”</li> </ul> Sobre a própria ciência: <ul style="list-style-type: none"> <li>• “Investigação científica” (meios)</li> <li>• “Explicação científica” (fins)</li> </ul>	Sobre o mundo natural: <ul style="list-style-type: none"> <li>• “Sistemas físicos”</li> <li>• “Sistemas vivos”</li> <li>• “Sistemas da terra e do espaço”</li> <li>• “Sistemas de tecnologia”</li> </ul> Sobre a própria ciência: <ul style="list-style-type: none"> <li>• “Investigação científica” (meios)</li> <li>• “Explicação científica” (fins)</li> </ul>
<b>Processos</b> (competências científicas)	<ul style="list-style-type: none"> <li>• Reconhecer questões investigáveis cientificamente</li> <li>• Identificar a evidência necessária para uma investigação científica</li> <li>• Tirar e avaliar conclusões</li> <li>• Comunicar conclusões válidas</li> <li>• Demonstrar compreensão de conceitos científicos</li> </ul>	<ul style="list-style-type: none"> <li>• Descrever, explicar e prever fenómenos científicos</li> <li>• Compreender a investigação científica</li> <li>• Interpretar evidências e conclusões científicas</li> </ul>	<ul style="list-style-type: none"> <li>• Identificar questões científicas</li> <li>• Explicar fenómenos cientificamente</li> <li>• Usar evidência científica</li> </ul>	<ul style="list-style-type: none"> <li>• Identificar questões científicas</li> <li>• Explicar fenómenos cientificamente</li> <li>• Usar evidência científica</li> </ul>
<b>Contextos</b> (situações que envolvem ciência e tecnologia)	<ul style="list-style-type: none"> <li>• Domínios de aplicação da ciência (com relevância pessoal, comunitária, global e histórica) agrupados em três grandes categorias:</li> <li>• Ciência, Terra e saúde</li> <li>• Ciência, Terra e ambiente</li> <li>• Ciência e Tecnologia</li> </ul>	<ul style="list-style-type: none"> <li>• O contexto da ciência, focando os usos em relação à:</li> <li>• Ciência, Terra e saúde</li> <li>• Ciência, Terra e ambiente</li> <li>• Ciência e Tecnologia</li> </ul>	<ul style="list-style-type: none"> <li>• Domínios de aplicação da ciência, focando os usos em contextos pessoais:</li> <li>• “Saúde”</li> <li>• “Recursos naturais”</li> <li>• “Desastres naturais”</li> <li>• “Fronteira da ciência e da tecnologia”</li> </ul>	<ul style="list-style-type: none"> <li>• Domínios de aplicação da ciência, focando os usos em contextos pessoais:</li> <li>• “Saúde”</li> <li>• “Recursos naturais”</li> <li>• “Desastres naturais”</li> <li>• “Fronteira da ciência e da tecnologia”</li> </ul>
<b>Atitudes</b> (para com a ciência)			<ul style="list-style-type: none"> <li>• Reações a questões científicas (ênfase nas duas primeiras):</li> <li>• Interesse pela ciência</li> <li>• Apoio à instigação científica</li> <li>• Responsabilidade face aos recursos naturais e do ambiente</li> </ul>	<ul style="list-style-type: none"> <li>• Interesse pela ciência</li> <li>• Apoio à investigação científica</li> <li>• Responsabilidade face aos recursos naturais e do ambiente</li> </ul>

Fonte: relatórios de enquadramento conceptual (*framework*) do PISA (OCDE, 1999, 2003b, 2006, 2009c) e relatórios nacionais do (GAVE, 2003, 2004, 2007).

Tabela 8 | Descrições dos níveis de proficiência em literacia científica (2009)

Nível	Percentagem de alunos da OCDE que atingem, no mínimo, o nível	Limites mínimos de pontos	O que os alunos são tipicamente capazes de fazer...
6	1.1 %	708	<p>Neste nível um aluno consegue identificar, explicar e aplicar conhecimentos científicos e conhecimentos sobre ciência num leque variado de situações complexas do dia-a-dia. Conseguem também relacionar informação de diferentes fontes para explicar um determinado fenómeno ou para dar resposta a um dado problema concreto. Consegue ainda demonstrar claramente um raciocínio científico avançado para a procura de soluções para situações científicas novas. Um aluno com este nível de desempenho consegue utilizar conhecimentos científicos e aplicá-los em decisões de nível pessoal, social ou até mesmo global.</p>
5	8.5 %	633	<p>Neste nível um aluno consegue identificar componentes científicas de um vasto leque de complexas situações reais, aplicar conceitos e conhecimentos de ciência nesses situações e é capaz de comparar, selecionar e avaliar adequadamente o recurso a evidência científica para dar resposta a tais situações. A este nível os alunos conseguem utilizar de forma correta capacidades de questionar, de relacionar conhecimentos e de criticar situações reais com que se deparem, conseguindo traçar explicações baseadas em evidência científica e argumentos baseados na sua análise crítica.</p>
4	29.1 %	559	<p>Um aluno com um nível 4 de desempenho consegue lidar eficazmente com situações e assuntos que possam implicar a necessidade de fazer inferência sobre um determinado conjunto de fatos científicos. Consegue selecionar e integrar explicações e/ou argumentos de várias disciplinas científicas e relacioná-las com aspetos reais do dia-a-dia. É ainda capaz de refletir sobre as suas ações e tomar decisões recorrendo a conhecimentos científicos adquiridos.</p>
3	57.7 %	484	<p>Tipicamente um aluno de nível 3 de desempenho é capaz de usar conceitos científicos e de fazer previsões ou providenciar explicações bem como de reconhecer questões que podem ser respondidas pela investigação científica e/ou identificar pormenores do que uma investigação científica envolve. É ainda capaz de selecionar informação relevante a partir de dados variados e de tirar conclusões ou de fazer a sua própria avaliação de uma determinada situação.</p>
2	82.0 %	409	<p>Um aluno com nível de desempenho 2 possui conhecimentos científicos suficientes para conseguir fornecer explicações de situações familiares e tirar conclusões através de uma investigação simples. É ainda capaz de fazer interpretações muito simples de resultados de processos de investigação ou de um determinado problema.</p>
1	95.0 %	335	<p>Um aluno com este nível de desempenho demonstra possuir um conhecimento científico muito limitado conseguindo aplicá-lo a situações muito simples e já suas conhecidas. Apenas consegue apresentar explicações científicas óbvias e resultante de evidência científica.</p>

Fonte: relatório nacional do PISA 2006 (GAVE, 2007: 65 – 66); PISA 2009 (OCDE, 2010a: 147-150).

Tabela 9 | Evolução das dimensões da literacia em matemática

	2000	2003	2006	2009
<b>Conteúdos</b> (conceitos matemáticos relevantes)	Quatro temas amplos, centrando a atenção apenas dois: <ul style="list-style-type: none"> <li>• Espaço e forma</li> <li>• Mudança e relações;</li> </ul>	<ul style="list-style-type: none"> <li>• Quantidade;</li> <li>• Espaço e forma;</li> <li>• Mudança e relações;</li> <li>• Incerteza</li> </ul>	<ul style="list-style-type: none"> <li>• Quantidade;</li> <li>• Espaço e forma;</li> <li>• Mudança e relações;</li> <li>• Incerteza;</li> </ul>	<ul style="list-style-type: none"> <li>• Quantidade;</li> <li>• Espaço e forma;</li> <li>• Mudança e relações;</li> <li>• Incerteza;</li> </ul>
<b>Processos</b> (clusters/ constelações de competências)	Três classes de competências; <ul style="list-style-type: none"> <li>• Reprodução, definição e cálculos;</li> <li>• Conexões e integração para resolução de problemas;</li> <li>• Reflexão (matematização, pensamento matemático, generalização e perspicácia)</li> </ul>	<ul style="list-style-type: none"> <li>• Reprodução (operações matemáticas simples);</li> <li>• Conexão (ligar ideias para resolver problemas de resolução direta);</li> <li>• Reflexão (pensamento matemático mais abrangente)</li> </ul>	<ul style="list-style-type: none"> <li>• Reprodução</li> <li>• Conexão;</li> <li>• Reflexão;</li> </ul>	<ul style="list-style-type: none"> <li>• Reprodução</li> <li>• Conexão;</li> <li>• Reflexão;</li> </ul>
<b>Contextos</b> (situações)	Cinco situações principais; <ul style="list-style-type: none"> <li>• Pessoal;</li> <li>• Educacional</li> <li>• Ocupacional</li> <li>• Comunidade local e sociedade;</li> <li>• Científica;</li> </ul>	<ul style="list-style-type: none"> <li>• Pessoal;</li> <li>• Educacional e ocupacional</li> <li>• Comunidade local e sociedade;</li> <li>• Científica;</li> </ul>	<ul style="list-style-type: none"> <li>• Pessoal;</li> <li>• Educacional e ocupacional</li> <li>• Pública;</li> <li>• Científica;</li> </ul>	<ul style="list-style-type: none"> <li>• Pessoal;</li> <li>• Educacional e ocupacional</li> <li>• Pública;</li> <li>• Científica;</li> </ul>

Fonte: relatórios de enquadramento conceptual (*framework*) do PISA (OCDE, 1999, 2003b, 2006, 2009c) e relatórios nacionais do (GAVE, 2002, 2004).

**Tabela 10| Descrições dos níveis de proficiência da literacia matemática (2009)**

Nível	Percentagem de alunos da		O que os alunos são tipicamente capazes de fazer...
	OCDE que atingem, no mínimo, o nível	Limites mínimos de pontos	
6	3.1%	669	No nível 6, os alunos são capazes de concetualizar, generalizar e utilizar informação, com base nas suas investigações e na modelação de situações problemáticas complexas. Conseguem estabelecer a ligação entre diferentes fontes de informação e diferentes representações e fazer transferências entre elas, com flexibilidade. Estes alunos são capazes de aplicar a perspicácia e a compreensão, a par do domínio de operações e relações matemáticas simbólicas e formais, no desenvolvimento de novas abordagem e estratégias face a situações novas. São capazes de formular e comunicar com exatidão as suas reflexões no que respeita às suas descobertas, interpretações, argumentos, bem como a adequação dos mesmos às situações originais.
5	12.7%	607	No nível 5, os alunos conseguem desenvolver e trabalhar com modelos de situações complexas, identificando constrangimentos e especificando hipóteses. São capazes de selecionar, comparar e avaliar estratégias adequadas de resolução de problemas, para lidarem com problemas complexos relacionados com estes modelos. Neste nível, os alunos são capazes de trabalhar estrategicamente, usando capacidades mentais e de raciocínio amplas e bem desenvolvidas, representações adequadamente ligadas, caracterizações simbólicas e formais e a perspicácia apropriada a estas situações. Conseguem refletir sobre as suas ações, formular e comunicar as suas interpretações e raciocínios.
4	31.6%	545	No nível 4, os alunos são capazes de trabalhar eficazmente com modelos explícitos para situações concretas complexas, as quais podem envolver constrangimentos ou exigir a formulação de hipóteses. Conseguem selecionar e integrar representações diferentes, inclusivamente simbólicas, ligando-as diretamente a aspetos de situações da vida real. Neste nível, os alunos são capazes de utilizar capacidades bem desenvolvidas e de raciocinar de modo flexível, com alguma perspicácia, nestes contextos. São capazes de construir e de comunicar explicações e argumentos, com base nos seus argumentos, interpretações e ações.
3	56.0%	482	No nível 3, os alunos são capazes de executar procedimentos descritos com clareza, incluindo os que requerem decisões sequenciais. Conseguem selecionar e aplicar estratégias simples de resolução de problemas. Neste nível, os alunos são capazes de interpretar e usar representações, com base em diferentes fontes de informação e de raciocinar diretamente a partir delas. Conseguem desenvolver comunicações curtas, que relatam os seus resultados, interpretações e raciocínios.
2	78.0%	420	No nível 2, os alunos são capazes de interpretar e reconhecer situações em contextos que não requerem mais do que inferência direta. São capazes de extrair informação relevante de uma única fonte e fazer uso de um único modelo de representação. Os alunos conseguem aplicar algoritmos, fórmulas, procedimentos ou convenções a um nível básico. São capazes de efetuar raciocínios diretos e de fazer interpretações literais dos resultados.
1	92.0%	358	No nível 1, os alunos são capazes de responder a questões que envolvem contextos familiares, em que toda a informação relevante está presente e as questões são claramente definidas. São capazes de identificar a informação e de executar procedimentos de rotina, de acordo com desenvolvimento parte diretamente dos estímulos dados.

Fonte: relatório nacional do PISA (GAVE, 2004: 13 – 15); (OCDE, 2010<sup>a</sup>: 130 – 13).





## Capítulo IV

### Modelos de Regressão Multinível

#### 4.1 | Modelos de regressão linear

Quando se pretende estudar a relação entre uma ou mais variáveis, variáveis explicativas, e uma variável de interesse, variável resposta, pode-se utilizar o modelo de regressão linear clássico,

$$Y = X\beta + \varepsilon,$$

onde

- \*  $Y$  é o vetor  $n \times 1$  da variável resposta;
- \*  $X$  é uma matriz de dimensão  $n \times (K+1)$  das variáveis explicativas associada ao vetor  $\beta = (\beta_0, \beta_1, \dots, \beta_K)^T$  de parâmetros desconhecidos;
- \*  $K$  corresponde ao número de variáveis explicativas;
- \*  $\varepsilon$  é um vetor de erros aleatórios com distribuição  $N(0, \sigma^2 I)$ .

Neste modelo os parâmetros são fixos e o parâmetro  $\beta_0$  representa a ordenada na origem, ou seja, é o valor esperado da variável resposta  $Y$  para  $X = 0$ . Enquanto  $\beta_K$  é o efeito que a variável explicativa tem no valor esperado da variável resposta, isto significa que por cada unidade adicional em  $X_K$  o valor esperado da variável  $Y$  varia  $\beta_K$  quando as restantes variáveis explicativas não se alteram.

Nos modelos de regressão clássica assume-se a independência dos erros. Em estudos educacionais, os dados encontram-se estruturados de forma hierárquica, ou seja, os alunos correspondem a unidades de nível mais baixo (nível 1), que por sua vez se agrupam em unidades de nível mais alto, ou escolas (nível 2). Desta forma, os erros aleatórios de alunos pertencentes à mesma escola não serão, em princípio, independentes se a escola exercer algum impacto nos resultados. Assim, quando o pressuposto de independência for violado e as unidades do mesmo nível estiverem correlacionadas pode recorrer-se ao uso dos modelos de regressão multinível.

## 4.2 | Modelos multinível

A investigação dos modelos multinível teve a sua origem na área da educação. Nomeadamente teve por base um estudo muito conhecido (Bennett 1976) que aplicou modelos de regressão linear para demonstrar que alunos de uma escola primária do Reino Unido sujeitos ao chamado método formal de ensino apresentavam maior progresso do que os alunos que não estavam sujeitos a este método de ensino, o chamado método progressivo.

Posteriormente, Aitkin et al. (1981) demonstrou que quando a análise considerava agrupar os alunos em turmas, as diferenças significativas desapareciam e os alunos sujeitos ao ensino formal não demonstravam maior progresso relativamente aos outros alunos. Este estudo é o primeiro exemplo importante de uma análise multinível de dados nas ciências sociais. Foi então que Aitkin et al. (1981) e Longford (1986), dois matemáticos ingleses, escreveram um artigo que veio revolucionar o mundo da investigação educativa: os modelos multinível. Neste, foi demonstrado que os modelos de regressão linear usados para estudar a forma como um conjunto de variáveis explicavam uma variável produto, só poderiam ser empregue num caso muito especial: quando as observações eram independentes Gelman e Hill (2007), Goldstein (2003), Heck e Thomas (2000) e Hox (1998). No entanto, a nossa realidade, na qual os estudantes são agrupados em turmas, diferentes turmas estão agrupadas em escolas e as diferentes escolas em distritos, regiões ou países, não é compatível com a imposição de independência das observações.

Segundo Murillo (2008), os modelos multinível representam um dos métodos de análise mais interessantes em investigação quantitativa desenvolvidos nos últimos anos. Na área da educação, por exemplo, analisam-se conjuntamente as variáveis respeitantes aos alunos e as variáveis respeitantes à escola.

Os modelos multinível representam uma extensão do modelo de regressão clássica quando as variáveis são dispostas em vários níveis de agregação. Esta técnica é um tipo de análise de regressão que, simultaneamente, tem em consideração múltiplos níveis de agregação.

Estes modelos, quando utilizados corretamente, permitem obter melhores estimativas dos coeficientes de regressão. De acordo com Guo e Zhao (2000), quando a estrutura hierárquica existente nos dados é ignorada, os modelos lineares de regressão clássica tendem a subestimar os erros-padrão, o que resulta em intervalos de confiança e testes de hipóteses incorretos podendo originar conclusões erradas.

#### 4.2.1 | Modelos de regressão linear multinível com dois níveis

Considere-se uma amostra de alunos agrupados em escolas. Os alunos são identificados pelo índice  $i$  e escolas são identificadas pelo índice  $j$ . O índice  $j$  varia de 1 a  $J$  e o índice  $i$  varia de 1 a  $n_j$ , sendo  $J$  o número total de escolas na amostra e  $n_j$  o número total de alunos que pertence à escola  $j$ .

A análise multinível tem em conta a correlação intra-classes, isto é, a correlação entre duas unidades do nível 1 (alunos) dentro de um mesmo grupo de nível 2 (escolas) e produz desvios padrão corrigidos ao tratar as unidades do nível 2 como amostras aleatórias de uma população. A utilização de modelos multinível com dois níveis permite ainda considerar a escola como um fator que potencialmente afeta os resultados escolares dos alunos.

De seguida sugerem-se as etapas de análise multinível segundo Raudenbush and Bryk (2002).

**Etapa 1:** Análise de variância com um fator aleatório para determinar se existe ou não efeito escola.

Este modelo é designado por modelo nulo (modelo mais simples) porque não envolve nenhuma variável explicativa em ambos os níveis.

A equação que representa este modelo é

$$Y_{ij} = \beta_{0j} + r_{ij}, i = 1, \dots, n_j, j = 1, \dots, J$$
$$\beta_{0j} = \mu_{00} + u_{0j},$$

ou equivalentemente

$$Y_{ij} = \mu_{00} + u_{0j} + r_{ij}, i = 1, \dots, n_j, j = 1, \dots, J$$

em que

$Y_{ij}$  representa a variável resposta do aluno  $i$  na escola  $j$ ,

$\beta_{0j}$  representa a média da variável resposta na  $j$ -ésima escola,

$\mu_{00}$  representa a média global da variável resposta,

$r_{ij}$  é o erro aleatório associado ao  $i$ -ésimo aluno da escola  $j$  e  $r_{ij} \sim N(0, \sigma_r^2)$  e  $r_{ij}$ 's são independentes,

$u_{0j}$  é o erro aleatório associado à escola  $j$  e  $u_{0j} \sim N(0, \sigma_{u_0}^2)$ ,  $\text{cov}(u_{0j}, u_{0j'}) = 0$  para  $j \neq j'$  e  $\text{cov}(u_{0j}, r_{ij}) = 0$ .

**Etapa 2:** Após concluir a existência do efeito escola na etapa 1, nesta etapa tenta-se estimar o efeito escola quando as variáveis respeitantes aos alunos (nível 1) e que condicionam os seus resultados são também consideradas. Desta forma, obtém-se um modelo que estima de forma mais fiável o efeito escola. Nesta etapa, há vários modelos possíveis a estudar, e a procura do modelo mais adequado passa por explorar as várias alternativas.

Neste trabalho apenas vamos considerar o modelo de efeitos fixos com variáveis explicativas ao nível do aluno e da escola.

### Modelo de efeitos fixos

Consideremos inicialmente o modelo que inclui  $K$  variáveis explicativas ao nível do aluno,  $X_{kij}$ , onde  $k$  é o índice associado à variável explicativa  $X$  ( $k=0, \dots, K$ ),  $i$  representa o aluno ( $i=1, \dots, n_j$ ) e  $j$  representa a escola ( $j=1, \dots, J$ ) e  $S$  variáveis explicativas ao nível da escola  $W_{sj}$ , onde  $S$  é o índice associado à variável escola  $W$ .

Neste modelo os termos independentes  $\beta_{0j}$  são variáveis aleatórias que variam de escola para escola e os declives  $\beta_k$ , que representam o efeito de cada variável  $X_{kij}$  na variável resposta são fixos. Neste caso considera-se que o impacto de cada variável explicativa na variável resposta não varia de escola para escola.

Matematicamente representa-se por

$$Y_{ij} = \beta_{0j} + \beta_1 X_{1ij} + \dots + \beta_K X_{Kij} + r_{ij}, i = 1, \dots, n_j, j = 1, \dots, J$$

$$\beta_{0j} = \mu_{00} + \gamma_1 W_{1j} + \dots + \gamma_S W_{Sj} + u_{0j}$$

ou equivalentemente

$$Y_{ij} = \mu_{00} + \gamma_1 W_{1j} + \dots + \gamma_S W_{Sj} + u_{0j} + \beta_1 X_{1ij} + \dots + \beta_K X_{Kij} + r_{ij}, i = 1, \dots, n_j, j = 1, \dots, J$$

em que

$Y_{ij}$  representa a variável resposta do aluno  $i$  da escola  $j$ ,

$\beta_{0j}$  representa a média da variável resposta na  $j$ -ésima escola,

$\mu_{00}$  representa a média global da variável resposta,

$W_{Sj}$  representa a  $S$  variável explicativa ao nível da escola  $j$ ,

$\gamma_S$  representa o efeito da variável explicativa  $W_{Sj}$  na variável resposta,

$r_{ij}$  é o erro aleatório associado  $i$ -ésimo aluno da escola  $j$ ,  $r_{ij} \sim N(0, \sigma_r^2)$  e  $r_{ij}'s$  são independentes,

$u_{0j}$  é o erro aleatório associado à escola  $j$  e  $u_{0j} \sim N(0, \sigma_{u_0}^2)$ ,  $u_{0j}'s$  são independentes e  $\text{cov}(u_{0j}, r_{ij}) = 0$ .

A estimativa do parâmetro  $\hat{\beta}_{0j}$ , está centrada em torno da média global,  $\mu_{00}$ , e varia relativamente a esta por um termo aleatório  $u_{0j}$ . A estimativa do parâmetro  $\hat{\beta}_k$  representa o efeito da variável explicativa  $X_K$  na variável resposta por cada unidade adicional daquela.

### 4.3 | Modelo de regressão logística multinível

Como referem Guo e Zhao (2000), há um grande interesse no estudo de problemas caracterizados por dados com uma estrutura hierárquica e com variáveis dependentes dicotômicas.

No caso em estudo o modelo de regressão logística multinível, com dois níveis, trata o aluno como a unidade de nível 1, identificado pelo índice  $i$ , e a escola como unidade de nível 2 identificada pelo índice  $j$ . Considera-se a existência  $J$  escolas (o índice  $j$  varia de 1 a  $J$ ) cada uma delas com  $n_j$  alunos (o índice  $i$  varia de  $i$  até  $n_j$ ).

Seja  $y_{ij}$  o valor da variável resposta do  $i$ -ésimo aluno da  $j$ -ésima escola em que  $y_{ij} = 1$ , se o aluno  $i$  da escola  $j$  se encontrar na categoria de melhor desempenho (=1) e  $y_{ij} = 0$  caso contrário.

Assim, a variável resposta  $Y_{ij}$  deste modelo é uma variável binária tomando os valores 0 ou 1 com probabilidade  $1 - \pi_{ij}$  e  $\pi_{ij}$ , respetivamente. Assume-se que  $Y_{ij}$  é uma variável aleatória com distribuição Bernoulli,

$$E(Y_{ij}) = \pi_{ij}$$

e

$$Var(Y_{ij}) = 1 - \pi_{ij}.$$

No nosso estudo define-se a probabilidade de um aluno  $i$  pertencer à categoria de melhor desempenho como a probabilidade de a variável  $Y_{ij}$  assumir o valor 1, isto é,  $\pi_{ij} = P(Y_{ij} = 1)$  em que a probabilidade de sucesso  $\pi_{ij}$  é modelada usando a função de ligação  $\log it$ ,

$$\log it(\pi_{ij}) = \ln\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right)$$

A razão de probabilidade dada por  $\frac{\pi_{ij}}{1 - \pi_{ij}}$  é denominada *odd ratio*.

### 4.3.1 | Modelo nulo

O primeiro modelo a apresentar é o modelo nulo. Este é o modelo mais simples porque não envolve nenhuma variável explicativa em ambos os níveis, isto é, é um modelo em que o preditor linear é constituído apenas por uma constante e toda a variação nos valores observados é a variação residual não explicada pelo modelo.

$$\begin{aligned}\log it(\pi_{ij}) &= \beta_{0j}, i = 1, \dots, n_j, j = 1, \dots, J \\ \beta_{0j} &= \gamma_{00} + u_{0j}, u_{0j} \sim N(0, \sigma_{u0}^2)\end{aligned}$$

ou equivalentemente

$$\log it(\pi_{ij}) = \gamma_{00} + u_{0j}, u_{0j} \sim N(0, \sigma_{u0}^2), i = 1, \dots, n_j, j = 1, \dots, J \quad (1)$$

No modelo de regressão multinível é usado o coeficiente de correlação intra-classe, ICC (*Intra-class correlation*), que indica a percentagem da variância total que é explicada pelo fator escola, isto é, mede o grau em que os alunos de uma mesma escola são mais semelhantes entre si do que os alunos em escolas diferentes.

O ICC é dado por,

$$ICC = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \left(\frac{\pi^2}{3}\right)}$$

onde  $\sigma_{u0}^2$  é a variância dos erros do nível superior (variância entre escolas) e  $\frac{\pi^2}{3}$  (variância da distribuição logística) é a variância dos erros do nível inferior (variâncias entre os alunos) (Goldestein e Rasbash, 2005).

### 4.3.2 | Modelo de efeitos fixos

Neste trabalho consideramos apenas o modelo de efeitos fixos. Começamos por definir o modelo que inclui apenas variáveis explicativas ao nível do aluno,  $X_{kij}$ , onde  $k$  é o índice associado à variável explicativa  $X$  ( $k=1, \dots, K$ ),  $i$  representa o aluno ( $i=1, \dots, n_j$ ) e  $j$  representa a escola ( $j=1, \dots, J$ ).

No modelo de efeitos fixos os termos independentes  $\beta_{0j}$  são variáveis aleatórias que variam de escola para escola e os declives  $\beta_k$ , que representam o efeito de cada variável  $X_{kij}$  na variável de resposta, são fixos. Neste caso considera-se que o impacto de cada variável explicativa na variável de resposta não varie de escola para escola.

Matematicamente representa-se por,

$$\begin{aligned} \log it(\pi_{ij}) &= \beta_{0j} + \beta_1 X_{1ij} + \dots + \beta_K X_{Kij}, i=1, \dots, n_j, j=1, \dots, J \\ \beta_{0j} &= \gamma_{00} + u_{0j}, u_{0j} \sim N\left(0, \sigma_{u0}^2\right), k=1, \dots, K \end{aligned}$$

ou equivalentemente

$$\log it(\pi_{ij}) = \gamma_{00} + \beta_1 X_{1ij} + \beta_K X_{Kij} + u_{0j}, u_{0j} \sim N\left(0, \sigma_{u0}^2\right) \quad (2)$$

Os modelos de efeitos fixos anteriormente expostos só apresentavam variáveis explicativas ao nível do aluno (nível 1). Contudo nestes modelos podem ser introduzidas variáveis explicativas ao nível da escola (nível 2).



Matematicamente representa-se por,

$$\log it(\pi_{ij}) = \beta_{0j} + \beta_1 X_{1ij} + \dots + \beta_K X_{Kij}, i = 1, \dots, n_j, j = 1, \dots, J$$

$$\beta_{0j} = \gamma_{00} + \sum_{s=1}^S \gamma_s W_{sj} + u_{0j}, u_{0j} \sim N\left(0, \sigma_{u0}^2\right)$$

ou equivalentemente

$$\log it(\pi_{ij}) = \gamma_{00} + \sum_{s=1}^S \gamma_s W_{sj} + u_{0j} + \beta_1 X_{1ij} + \dots + \beta_K X_{Kij}, i = 1, \dots, n_j, j = 1, \dots, J \quad (3)$$

## 4.4| Métodos de Estimação

Para os modelos de regressão logística multinível existem dois métodos de estimação dos parâmetros de interesse: o método o método de quasi-verosimilhança ou aproximações numéricas aos estimadores de máxima verosimilhança. Neste ultimo caso é comum recorrer-se ao método de integração numérica ou a quadratura Gauss-Hermite.

### 4.4.1| Método – A quasi-verosimilhança (QVP)

Neste método a linearização da função de ligação é feita pela expansão da série de Taylor, podendo ser de primeira ou segunda ordem. Caso a expressão resultante da integração dos efeitos aleatórios não seja analiticamente possível, pode-se recorrer à integração numérica com quadratura Gaussiana. Alguns procedimentos, no entanto, podem evitar um grande custo computacional, como, por exemplo, a quasi-verosimilhança preditiva ou penalizada (PQL), a aproximação de Laplace e a quasi-verosimilhança marginal (MQL).

#### 4.4.2| Método – A integração numérica ou quadratura de Gauss-Hermite (QGH)

Neste segundo método de estimação, a verossimilhança analiticamente intratável pode ser solucionada numericamente através da quadratura Gaussiana ordinária (GQ) ou a sua versão melhorada, a quadratura Gaussiana Adaptativa (AGQ) (Rabe-Hesketh, 2005). Os métodos de integração numérica maximizam a verossimilhança correta e, neste caso, testes e índices para avaliação da qualidade do ajustamento baseados no desvio podem ser usados (Hox, 2002).

### 4.5| Testes Estatísticos

#### 4.5.1| Teste de Wald

Após a estimação dos parâmetros do modelo de regressão multinível deve-se analisar a significância estatística de cada um dos parâmetros estimados.

Para realizar testes de hipóteses sob o vetor dos coeficientes do modelo de regressão multinível pode-se aplicar o teste de Wald. Este teste é utilizado para verificar a significância estatística dos parâmetros fixos do modelo. Para testar a hipótese

$$H_0 : \beta_j = 0 \quad \text{vs} \quad H_1 = \beta_j \neq 0,$$

com  $j = 1, \dots, J$ , a estatística de Wald, que sob a hipótese nula ( $H_0$ ) segue uma distribuição assintótica gaussiana é dada por,

$$W = \frac{\hat{\beta}_j}{\sqrt{\text{var}(\hat{\beta}_j)}} \sim N(0,1)$$

Nos modelos de regressão multinível, podem ainda ser realizados testes de hipóteses em relação aos parâmetros de variância após determinar as estimativas da variância e o respetivo erro-padrão pode-se aplicar o teste de Wald (Hox, 2002). Contudo, Raudenbush e Bryk (2002) e Hox (2002) referem que esse teste, quando aplicado às variâncias, deve ser feito unilateralmente, já que é uma forma conservadora de considerar que o pressuposto da aproximação à distribuição Normal pode não ser o mais adequado. Para além disso não faz sentido considerar a hipótese alternativa de a variância assumir um valor negativo.

#### 4.6| Qualidade de ajustamento

A qualidade de ajustamento no modelo de regressão multinível pode ser avaliada recorrendo a várias estatísticas (Hox, 2002).

O coeficiente de determinação proposto por McFadden é dado por,

$$R^2_{MF} = 1 - \frac{LL(M)}{LL(M_0)}$$

onde  $LL(M)$  é -2 vezes o logaritmo da função de verosimilhança do modelo estimado e  $LL(M_0)$  é -2 vezes o logaritmo da função verosimilhança do modelo nulo.

Um outro coeficiente de determinação que considera a dimensão da amostra é o coeficiente proposto por Cox e Snell dado por

$$R^2_{CS} = 1 - \exp\left(\frac{LL(M) - LL(M_0)}{n}\right)$$

O coeficiente de ajustamento proposto por Nagelkerke tem a vantagem de assumir o valor 1 como valor máximo,

$$R^2_N = \frac{R^2_{CS}}{1 - \exp\left(\frac{-LL(M_0)}{n}\right)}$$

Embora estes coeficientes não possam ser interpretados de forma independente ou comparados em diferentes conjuntos de dados, eles são úteis para comparar e avaliar diferentes modelos de previsão com a mesma variável resposta no mesmo conjunto de dados.

Em geral, estes coeficientes de determinação assumem valores baixos e valores que variam entre 0.2 e 0.4 indicam um bom ajustamento.

## Capítulo V

### Análise Exploratória

#### 5.1 | Dados e Variáveis

A base de dados utilizada na realização deste trabalho foi a base de dados PISA 2009, recolhida em Dezembro de 2011. Os instrumentos de recolha de informação desenvolvidos e aplicados no âmbito do projeto PISA são variados (como foram mencionados no Capítulo II).

A base de dados do PISA é constituído por três bases de dados: dos Estudante, das Escolas e dos Pais/Encarregados de Educação, de onde foram retiradas as variáveis consideradas importantes para o nosso estudo. Algumas das variáveis disponíveis correspondem a respostas concretas aos questionários realizados aos alunos, às escolas e aos encarregados de educação. Outras resultam da aplicação da Análise em Componentes – método estatístico que transforma um conjunto de variáveis que se encontram correlacionadas num novo conjunto de variáveis não correlacionadas designadas por componentes principais e que resultam de combinações lineares do conjunto inicial.

As variáveis selecionadas apresentavam uma codificação específica do PISA, mas, para uma mais fácil compreensão e tratamento procedeu-se à sua devida recodificação.

Relativamente às variáveis explicativas utilizaram-se variáveis ao nível do aluno e ao nível da escola. Estas foram escolhidas tendo em conta as variáveis usadas em outros estudos semelhantes.

De seguida descrevem-se sucintamente as variáveis em estudo assim como a nomenclatura e a categorização usada.

### Variáveis do Aluno

#### Demográficas

- \* **Sexo do aluno (*SEX*)** – representa o sexo do aluno:  
1 – Feminino; 2 – Masculino.
  
- \* **Idade do aluno (*AGE*)** – representa a idade do aluno.
  
- \* **Situação de imigração (*IMI*)** – indica se o aluno é imigrante ou não:  
1 – Nativo; 2 – Imigrante.
  
- \* **Nível de ensino que frequenta (*NEF*)** – representa o nível de ensino que o aluno frequenta:  
1 – 3º Ciclo do Ensino básico; 2 – Ensino Secundário - curso tecnológico ou profissional/ Formação Profissional; 3 – Ensino Secundário - curso geral/ científico – humanístico.
  
- \* **Região (*REG1*)** – indica a região a que cada aluno pertence:  
1 – Alentejo; 2 – Algarve; 3 – Centro; 4 – Lisboa; 5 – Norte; 6 – Região Autónoma da Madeira; 7 – Região Autónoma dos Açores; 8 – Desconhecido.
  
- \* **Expectativas do aluno (*EXPAL*)** – indica o grau de ensino que o aluno espera completar:  
1 – 3º Ciclo do Ensino básico; 2 – Ensino secundário - curso tecnológico ou profissional/ Formação Profissional; 3 – Ensino secundário - curso geral/científico - humanístico;  
4 – Licenciatura (ou equivalente).

#### Familiares

- \* **Constituição do agregado familiar (*AGREG*)** – indica o agregado familiar do aluno:  
1 – Um parente; 2 – Família tradicional; 3 – Outra estrutura; 4 – Desconhecido.

- \* **Ajuda nos trabalhos de casa (TPC)** – indica se os pais ajudam nos trabalhos de casa:  
1 – Nunca; 2 – Uma a duas vezes por mês; 3 – Uma a duas vezes por semana; 4 – Todos os dias; 5 – Desconhecido.
  
- \* **Nº médio de horas de explicação a ciências (HEXPCN)** – representa o número de horas de explicação por semana a ciências. Esta variável foi criada com base na variável *horas por semana em aulas suplementares/ apoio/ explicações*, da base de dados inicial:  
1 – Não tem; 2 – Tem entre 0 a 2 horas por semana; 3 – Tem entre 2 a 4 horas por semana; 4 – Tem mais de 4 horas por semana.
  
- \* **Nº médio de horas de explicação a matemática (HEXPMAT)** – representa o número de horas de explicação por semana a matemática. Esta variável foi criada com base na variável *horas por semana em aulas suplementares/ apoio/ explicações*, da base de dados inicial:  
1 – Não tem; 2 – Tem entre 0 a 2 horas por semana; 3 – Tem entre 2 a 4 horas por semana; 4 – Tem mais de 4 horas por semana.



- \* **Índice do uso de estratégias de controlo (ESTC)** – as estratégias de controlo são um fator construído com base em respostas dos alunos sobre:
  - a forma como os alunos estudam;
  - a capacidade de descobrir o que eles precisam aprender quando estudam;
  - a capacidade de verificar se eles entendem o que leem;
  - a capacidade de descobrir os conceitos que eles ainda não entenderam;
  - a capacidade para se certificar de que eles se lembram das questões mais importantes no que leram;
  - a procura de informações adicionais para esclarecer conceitos não compreendidos.

- \* **Índice do uso de estratégias de elaboração (ESTE)** – as estratégias de elaboração são um fator construído com base em respostas dos alunos sobre:
  - a relação entre novas informações e conhecimentos prévios;
  - o uso de informações aprendidas na escola e fora da escola;
  - a relação entre materiais aprendidos e experiências pessoais.
  
- \* **Índice do uso de estratégias de sumarização (ESTS)** – as estratégias de sumarização são um fator construído com base em respostas dos alunos sobre:
  - verificar cuidadosamente se os factos mais importantes do texto estão representados no resumo;
  - ler o texto, sublinhando as frases mais importantes.
  
- \* **Índice do uso de estratégias de compreender e relembrar (ESTCR)** – as estratégias de compreender e relembrar são um fator construído com base em respostas dos alunos sobre:
  - depois de ler o texto, discutir o seu conteúdo com outras pessoas;
  - sublinhar as partes importantes do texto;
  - resumir o texto com as próprias palavras;
  - ler rapidamente o texto duas vezes;
  - ler o texto em voz alta para outra pessoa.
  
- \* **Índice do uso de estratégia de memorização (ESTM)** – é um fator construído a partir da frequência com que os alunos utilizaram as seguintes estratégias ao estudar:
  - memorizar tudo o que possa ser abrangido;
  - memorizar tanto quanto possível;
  - memorizar todo o material novo para que eu possa pronunciar;
  - memorizar, ler em voz alta.



- \* **Aluno repetente (REP)** – indica se o aluno já repetiu pelo menos uma vez um nível de ensino:  
1 – Não repetente; 2 – Repetente.
  
- \* **Índice económico, social e cultural (ESCS)** – representa o índice socioeconómico e cultural do aluno. É um fator construído a partir do:
  - índice de bem-estar: traduzindo a disponibilidade de recursos em casa, por exemplo, a existência de eletrodomésticos, *software* educacional, presença de televisão, número de assoalhadas da casa, número de telemóveis e a posse de automóvel;
  - comunicação cultural dos pais: exprime a frequência com que os pais se comprometem com a discussão de assuntos políticos e sociais em casa, a troca de ideias sobre livros, filmes ou programas televisivos e com a audição de música clássica;
  - posses culturais da família: relacionadas com a existência em casa de itens tais como literatura clássica, livros de poesia e trabalhos de arte e com a situação ocupacional mais elevada dos pais (em anos escolares).

#### Variáveis da Escola



- \* **Tipo de escola (TIPO)** – representa o tipo de escola que cada aluno frequenta:  
1 – Escola Pública; 2 – Escola Privada dependente do governo; 3 – Escola Privada independente do governo.
  
- \* **Localização da escola (LOC)** – representa a localização onde a escola se encontra:  
1 – Aldeia; 2 – Cidade.
  
- \* **Dimensão da escola (DIM)** – representa a dimensão da escola em número de alunos.  
Esta variável foi codificada do seguinte modo:  
1 – Pequena [73, 850[; 2 – Média [850, 1626[; 3 – Grande [1626, 2401].

- \* **Número de ciclo da escola (CICLOS)** – representa os níveis de ensino da escola:  
1 – Todos os níveis; 2 – 1º2º3ºCiclos; 3 – 2º3ºCiclos, Secundário; 4 – 3ºCiclo, Secundário;  
5 – Secundário; 6 – 2º3ºCiclos; 7 – Desconhecido.
- \* **Média do índice económico, social e cultural (ESCSMED)** – é a média da variável *ESCS* da escola.
- \* **Percentagem de raparigas (PRAP)** – representa o rácio entre o número de raparigas da escola e o número total de alunos da escola.

**Recursos**

- \* **Rácio aluno/professores (RAP)** – representa a razão entre o número total de alunos inscritos e o número total de professores na escola (os professores com horário completo contribuem com um peso 1.0 e os professores com horário incompleto com um peso 0.5 no cálculo do número total de professores).
- \* **Rácio computador/aluno (RPCA)** – resulta do quociente entre o número de computadores disponíveis para fins educacionais e o número total de alunos da escola.
- \* **Proporção de computadores ligados à Internet (PCNET)** – representa a razão entre o número de computadores ligados à internet e o número de computadores disponíveis para fins educacionais.

**Políticas/ Estratégias**

- \* **Falta de professores de ciências (FPCN)** – representa a falta de professores de ciências na escola:  
1 – Não; 2 – Sim; 3 – Desconhecido.
  
- \* **Falta de professores de matemática (FPMAT)** – representa a falta de professores de matemática na escola:  
1 – Não; 2 – Sim; 3 – Desconhecido.
  
- \* **Comportamento do aluno (COMPA)** – é um fator relacionado com o absentismo escolar, a perturbação e falta às aulas, a falta de respeito, o uso de álcool ou drogas e a intimidação e provocação de outros alunos. Um maior valor nesta variável significa um pior comportamento do aluno.
  
- \* **Comportamento do professor (COMPP)** – é um fator relacionado com as expectativas dos professores, as relações aluno/professor, o conhecimento dos professores das necessidades individuais dos alunos, o absentismo dos professores, a resistência à mudança e o estímulo para o rendimento.
  
- \* **Índice de seletividade (SEL)** – representa a admissão dos alunos nas escolas. Pode basear-se em recomendações da escola de origem ou no registo académico. Os diretores são questionados relativamente à medida que estes fatores são considerados aquando a admissão:  
1 – Nunca; 2 – Às vezes; 3 – Sempre.
  
- \* **Criação de turmas de nível (CTN)** – refere-se à existência de turmas de nível na escola:  
1 – Para todos os assuntos; 2 – Para alguns assuntos; 3 – Para nenhum assunto;  
4 – Desconhecido.

- \* **Índice de atividades extracurriculares (AEC)** – é um fator que resulta da combinação das várias atividades extracurriculares oferecidas pela escola.

<b>Variáveis Resposta</b>
---------------------------

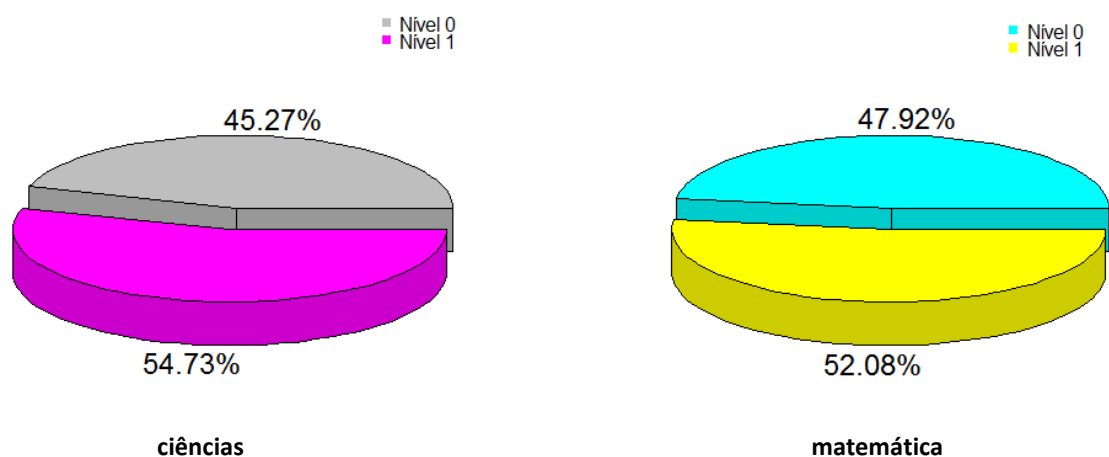
- \* **Nível de desempenho a ciências (CN)** – representa o nível de desempenho dos alunos a ciências. O nível de desempenho foi categorizado em dois níveis:
  - 0 – corresponde aos níveis de desempenho 1, 2 e 3 obtidos nos testes de ciências do PISA;
  - 1 – corresponde aos níveis de desempenho 4, 5 e 6 obtidos nos testes de ciências do PISA.
- \* **Nível de proficiência a matemática (MAT)** – representa o nível de desempenho dos alunos a matemática. O nível de desempenho foi categorizado em dois níveis:
  - 0 – corresponde aos níveis de desempenho 1, 2 e 3 obtidos nos testes de matemática do PISA;
  - 1 – corresponde aos níveis de desempenho 4, 5 e 6 obtidos nos testes de matemática do PISA.

## 5.2 | Análise Exploratória

Nesta secção vamos apresentar uma análise exploratória dos dados, com o objetivo de estudar cada uma das variáveis assim como a sua relação com as variáveis respostas *MAT* e *CN*.

Para colmatar os problemas da existência dos valores em falta (*missing values*) na base de dados, em cada variável explicativa contínua ao nível dos alunos (ESCS, ESTC, ESTE, ESTS, ESTM, ESTCR), os valores desconhecidos foram substituídos pela média da escola e em cada

variável explicativa ao nível da escola (RAP, RPCA, COMPA, COMPP, AEC, PCNET), os mesmos valores foram substituídos pela média nacional da variável explicativa correspondente. Este procedimento foi adotado no relatório técnico do PISA (OCDE, 2007).



**Figura 1 | Percentagem de alunos nos dois níveis de desempenho às disciplinas de ciências e matemática**

Pela análise da Figura 1 podemos concluir que a maioria dos alunos nas disciplinas de ciências e matemática se encontram em níveis de desempenho mais elevados.

As estatísticas descritivas (média, desvio padrão, mínimo, máximo, coeficiente de assimetria e coeficiente de achatamento) apresentadas na Tabela 11 permitem dar uma visão geral das variáveis explicativas contínuas.

Relativamente ao coeficiente de assimetria pode-se afirmar que as variáveis AGE, ESCS, ESCSMED, PRAP, RAP, RPCA, COMPA e COMPP apresentam valores positivos logo, possuem assimetria à direita, enquanto as variáveis ESTC, ESTE, ESTCR, ESTM, PCNET e AEC apresentam valores negativos, portanto possuem assimetria à esquerda.

As variáveis AGE, ESTS, ESTCR e ESCS apresentam caudas leves (valores do coeficiente de achatamento menores que 3) e as variáveis ESTC, ESTE, ESTM, PCNET, ESCSMED, PRAP, RAP, RPCA, COMPA, COMPP e AEC apresentam caudas pesadas.

Tabela 11 | Estatísticas descritivas das variáveis explicativas contínuas

<i>Nível</i>	<i>Variável</i>	<i>nº</i> <i>observações</i>	<i>Média</i>	<i>Desvio</i> <i>Padrão</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Coefficiente</i> <i>de</i> <i>Assimetria</i>	<i>Coefficiente</i> <i>de</i> <i>Achatamento</i>
Aluno	<i>AGE</i>	6298	15.290	0.453	15.250	16.250	0.939	1.881
	<i>ESTC</i>	6298	0.114	0.984	-3.453	2.504	-0.085	4.520
	<i>ESTE</i>	6298	0.387	0.848	-2.410	2.757	-0.104	4.667
	<i>ESTS</i>	6298	-0.008	1.034	-2.010	1.344	-0.605	2.400
	<i>ESTCR</i>	6298	-0.087	1.057	-1.882	1.415	-0.216	1.949
	<i>ESTM</i>	6298	-0.273	0.965	-3.020	2.690	-0.200	4.203
	<i>ESCS</i>	6298	-0.301	1.165	-3.132	3.224	0.404	2.389
Escola	<i>PCNET</i>	214	0.944	0.117	0.330	1.00	-2.809	11.364
	<i>ESCSMED</i>	214	-0.301	0.621	-1.560	1.610	0.772	3.696
	<i>PRAP</i>	214	0.506	5.414	26.030	89.220	1.066	12.642
	<i>RAP</i>	214	0.082	2.281	2.410	16.880	1.029	5.202
	<i>RPCA</i>	214	0.552	0.297	0.180	2.290	1.927	8.900
	<i>COMPA</i>	214	-0.053	0.909	-2.120	2.360	0.478	3.411
	<i>COMPP</i>	214	0.090	0.861	-2.250	2.120	0.404	3.223
	<i>AEC</i>	214	0.262	0.855	-3.260	2.910	-0.369	5.132

As Tabelas 12 e 13 resumem as variáveis explicativas categóricas indicando as diferentes categorias que cada variável pode assumir e a frequência absoluta e relativa de alunos/escola que estão em cada categoria.

Tabela 12 | Distribuição dos alunos

	<i>Ao nível do aluno</i>	
	Frequência absoluta	Frequência relativa
<b>Sexo (SEX)</b>		
Feminino	3278	52.05%
Masculino	3020	47.95%
<b>Situação de imigrante (IMI)</b>		
Nativo	5867	93.16%
Imigrante	431	6.84%
<b>Nível de ensino que frequenta (NEF)</b>		
3º Ciclo do Ensino básico	2211	35.40%
Ensino Secundário – curso geral/científico-humanístico	3118	49.50%
Ensino Secundário – curso tecnológico ou profissional	951	15.10%
<b>Região (REG)</b>		
Alentejo	832	13.21%
Algarve	611	9.70%
Centro	1400	22.23%
Lisboa	1161	18.43%
Norte	1922	30.52%
R.A.Açores	91	1.45%
R.A.Madeira	102	1.62%
Desconhecido	179	2.84%
<b>Expectativas do aluno (EXPAL)</b>		
3º Ciclo do Ensino básico	141	2.24%
Ensino secundário – curso tecnológico ou profissional/ Formação Profissional	1897	30.12%
Ensino secundário – curso geral/científico – humanístico	970	15.40%
Licenciatura (ou equivalente)	3290	52.24%
<b>Agregado familiar (AGREG)</b>		
Família tradicional	5011	79.56%
Um parente	1036	16.45%
Outros	182	2.89%
Desconhecido	69	1.10%
<b>Ajuda nos T.P.C. (TPC)</b>		
Nunca	1223	19.42%
Uma a duas vezes por mês	1984	31.50%
Uma a duas vezes por semana	2198	34.90%
Todos os dias	831	13.20%
Desconhecido	62	0.98%
<b>Horas de explicação a matemática (HEXPMAT)</b>		
Não tem	3226	51.22%
Tem entre 0 a 2 horas por semana	1214	19.28%
Tem entre 2 a 4 horas por semana	715	11.35%
Tem mais de 4 horas por semana	1143	18.15%
<b>Horas de explicação a ciências (HEXPCN)</b>		
Não tem	4353	69.12%
Tem entre 0 a 2 horas por semana	392	6.22%
Tem entre 2 a 4 horas por semana	196	3.11%
Tem mais de 4 horas por semana	1357	21.55%
<b>Repetente (REP)</b>		
Não repetente	3817	60.60%

Repetente	2481	39.40%
-----------	------	--------

Tabela 13 | Distribuição das escolas

	Ao nível da escola	
	Frequência absoluta	Frequência relativa
<b>Tipo de escola (TIPO)</b>		
Escola Pública	185	86.45%
Escola Privada dependente do governo	18	8.24%
Escola Privada independente do governo	11	5.31%
<b>Localização (LOC)</b>		
Aldeia	80	37.38%
Cidade	134	62.62%
<b>Dimensão da escola (DIM)</b>		
Pequena	119	55.61%
Média	83	38.78%
Grande	12	5.61%
<b>Número de ciclos da escola (CICLOS)</b>		
Todos os níveis	15	7.01%
1º2º3ºCiclos	25	11.68%
2º3ºCiclos, Secundário	23	10.75%
3ºCiclo, Secundário	81	37.85%
Secundário	29	13.55%
2º3ºCiclos	39	18.22%
Desconhecido	2	0.94%
<b>Falta de professores de matemática (FPMAT)</b>		
Não	204	95.33%
Sim	7	3.27%
Desconhecido	3	1.40%
<b>Falta de professores de ciências (FPCN)</b>		
Não	203	94.86%
Sim	8	3.74%
Desconhecido	3	1.40%
<b>Índice de seletividade (SEL)</b>		
Nunca	152	71.03%
Às vezes	57	26.63%
Sempre	5	2.34%
<b>Criação de turmas de nível (CTN)</b>		
Para todos os assuntos	139	64.95%
Para alguns assuntos	51	23.83%
Para nenhum assunto	20	9.35%
Desconhecido	4	1.87%

No nosso estudo, 52.05% dos estudantes são do sexo feminino, somente 6.84% dos estudantes são imigrantes, 60.60% não são repetentes, 30.52% dos alunos pertencem à região Norte e é nas regiões autónomas (Madeira e Açores) que se encontram o menor número de alunos, 1.62% e 1.45%, respetivamente e 79.56% dos alunos possui uma família



tradicional. Relativamente às expectativas dos alunos, 3290 (52.24%) alunos esperam obter uma Licenciatura (ou equivalente) e apenas 141 (2.24%) alunos esperam terminar os estudos no 3º Ciclo do Ensino básico.

Neste estudo, 69.12% dos alunos não frequentam explicações a ciências e 21.55% tem mais de 4 horas de explicação por semana a ciências. Relativa à disciplina de matemática, 51.22% dos alunos não tem explicações e 18.15% tem mais de 4 horas de explicação por semana.

Quanto às ajudas nos trabalhos de casa pelos pais, 19.42% dos estudantes não tem ajuda, 35% tem ajuda uma a duas vezes por semana e somente 13.20% tem todos os dias.

Observa-se ainda que, em aproximadamente 95% das escolas não há falta de professores nem de ciências nem de matemática. As escolas encontram-se na sua maioria em cidades (62.62%) e são de pequenas dimensões (55.61%).

As escolas públicas são 185, as escolas privadas dependentes do governo são 18 e as escolas privadas independentes do governo são 11. 37.85% das escolas possuem 3º Ciclo e Secundário, 64.95% criam de turmas de nível para todos os assuntos e em 152 escolas, a admissão dos alunos nunca se baseia em recomendações da escola de origem ou no registo académico.

Na Tabela 14 mostra-se como os alunos dos dois níveis de desempenho se distribuem de acordo com as suas características.

Os alunos do sexo feminino encontram-se em maior número no nível de desempenho 1 à disciplina de ciências, enquanto na disciplina de matemática são os alunos do sexo masculino que se encontram em maior número no nível de desempenho 1.

Em ambas as disciplinas, os alunos que frequentam o 3º Ciclo do ensino básico encontram-se no nível de desempenho 0 enquanto os alunos que frequentam o Ensino secundário - curso geral/científico humanístico que se encontram no nível de desempenho mais elevado.

Os alunos que não têm explicações quer a ciências quer a matemática, encontram-se maioritariamente no nível de desempenho 1, em ambas as disciplinas.

Também é de destacar que, em ambas as disciplinas, os alunos não repetentes encontram-se na sua grande maioria no nível de desempenho 1.

Tabela 14 | Tabela de contingência entre as variáveis resposta CN e MAT e as variáveis explicativas qualitativas ao nível do aluno

	Variáveis explicativas	CN		MAT	
		Nível 0	Nível 1	Nível 0	Nível 1
<b>SEX</b>	Feminino	1437	1841	1648	1630
	Masculino	1414	1606	1370	1650
<b>IMI</b>	Nativo	2577	3290	2748	3119
	Imigrante	274	157	270	161
<b>NEF</b>	3º Ciclo do ensino básico	1592	637	1680	549
	Ensino secundário – curso geral/científico humanístico	579	2539	680	2438
	Ensino secundário – curso tecnológico ou profissional	680	271	658	293
<b>EXPAL</b>	3º Ciclo do ensino básico	98	43	94	47
	Ensino secundário – curso geral/científico humanístico	558	412	589	381
	Ensino secundário – curso tecnológico ou profissional	1278	619	1305	592
	Licenciatura (ou equivalente)	917	2373	1030	2260
<b>AGREG</b>	Família tradicional	2170	2841	2270	2741
	Um parente	483	553	538	498
	Outros	137	45	146	36
	Desconhecido	61	8	64	5
<b>TPC</b>	Nunca	570	653	592	631
	Uma a duas vezes por mês	804	1180	852	1132
	Uma a duas vezes por semana	967	1231	1035	1163
	Todos os dias	458	373	487	344
	Desconhecido	52	10	52	10
<b>HEXPCN</b>	Não tem	1697	2656	1795	2558
	Tem entre 0 a 2 horas por semana	147	245	159	233
	Tem entre 2 a 4 horas por semana	81	115	82	114
	Tem mais de 4 horas por semana	926	431	982	375
<b>HEXPMAT</b>	Não tem	1250	1976	1316	1910
	Tem entre 0 a 2 horas por semana	543	671	593	621
	Tem entre 2 a 4 horas por semana	259	456	281	434
	Tem mais de 4 horas por semana	799	344	828	315
<b>REP</b>	Não repetente	919	2898	994	2823
	Repetente	1932	549	2024	457

Para averiguar possíveis associações entre as variáveis explicativas categóricas e a variável resposta é possível realizar o teste do qui – quadrado  $\chi^2$ . As hipóteses a serem testadas são:

$H_0$ : As variáveis são independentes

vs

$H_1$ : As variáveis são dependentes

**Tabela 15 | Teste de independência entre as variáveis resposta CN e MAT e as variáveis explicativas qualitativas ao nível do aluno**

<i>Variáveis explicativas categóricas</i>	<i>CN</i>		<i>MAT</i>	
	$\chi^2$	p - value	$\chi^2$	P – value
<i>SEX</i>	5.527	0.019	14.999	0.000
<i>IMI</i>	61.778	0.000	39.566	0.000
<i>NEF</i>	1657.360	0.000	1740.463	0.000
<i>REG</i>	105.462	0.000	119.910	0.000
<i>EXPAL</i>	931.777	0.000	808.597	0.000
<i>AGREG</i>	126.528	0.000	152.112	0.000
<i>TPC</i>	90.152	0.000	90.531	0.000
<i>HEXPCN</i>	381.965	0.000	-----	-----
<i>HEXPMAT</i>	-----	-----	423.539	0.000
<i>REP</i>	1754.143	0.000	1856.215	0.000

Pela análise da Tabela 15, no caso das variáveis dependentes CN e MAT, todas as variáveis explicativas categóricas ao nível do aluno parecem apresentar evidência para rejeitar a hipótese de independência, a um nível de significância de 5%.

Nas Tabelas 16 e 17 apresentam-se as estatísticas descritivas das variáveis explicativas contínuas ao nível do aluno para cada um dos níveis da variável resposta.

Tabela 16 | Estatísticas descritivas das variáveis explicativas contínuas por nível da variável resposta CN

Variáveis explicativas	CN	nº observações	Média	Desvio Padrão	Mínimo	Máximo
<b>AGE</b>	nível 0	2851	15.280	0.451	15.000	16.000
	nível 1	3447	15.290	0.454	15.000	16.000
<b>ESTC</b>	nível 0	2851	-0.232	0.956	-3.453	2.504
	nível 1	3447	0.400	0.912	-3.453	2.504
<b>ESTE</b>	nível 0	2851	0.214	0.815	-2.410	2.757
	nível 1	3447	0.530	0.848	-2.410	2.757
<b>ESTS</b>	nível 0	2851	-0.482	1.064	-2.010	1.344
	nível 1	3447	0.383	0.824	-2.010	1.344
<b>ESTCR</b>	nível 0	2851	-0.531	1.021	-1.882	1.415
	nível 1	3447	0.279	0.939	-1.882	1.415
<b>ESTM</b>	nível 0	2851	-0.200	0.941	-3.020	2.690
	nível 1	3447	-0.333	0.980	-3.020	2.690
<b>ESCS</b>	nível 0	2851	-0.741	0.988	-3.132	3.198
	nível 1	3447	0.064	1.173	-2.920	3.224

Tabela 17 | Estatísticas descritivas das variáveis explicativas contínuas por nível da variável resposta MAT

Variáveis explicativas	MAT	nº observações	Média	Desvio Padrão	Mínimo	Máximo
<b>AGE</b>	nível 0	3018	15.29	0.452	15.00	16.00
	nível 1	3280	15.29	0.454	15.00	16.00
<b>ESTC</b>	nível 0	3018	-0.174	0.966	-3.453	2.504
	nível 1	3280	0.379	0.924	-3.453	2.504
<b>ESTE</b>	nível 0	3018	0.225	0.811	-2.410	2.757
	nível 1	3280	0.536	0.854	-2.410	2.757
<b>ESTS</b>	nível 0	3018	-0.422	1.068	-2.010	1.344
	nível 1	3280	0.373	0.837	-2.010	1.344
<b>ESTCR</b>	nível 0	3018	-0.492	1.021	-1.882	1.415
	nível 1	3280	0.285	0.939	-1.882	1.415
<b>ESTM</b>	nível 0	3018	-0.201	0.946	-3.020	2.690
	nível 1	3280	-0.339	0.977	-3.020	2.690
<b>ESCS</b>	nível 0	3018	-0.745	0.988	-3.132	3.198
	nível 1	3280	0.109	1.165	-2.766	3.224

Esta tabela sugere que a idade média dos alunos que se encontram no nível de desempenho 1 não difere das médias dos alunos do nível e desempenho 0 em ambas as disciplinas.

Como a amostra é de grandes dimensões (6298 alunos) e as populações são independentes vai ser aplicado o teste t para testar a igualdade das médias das variáveis explicativas contínuas entre os alunos do nível 0 ( $\mu_0$ ) e os alunos do nível 1 ( $\mu_1$ ):

$$H_0 : \mu_0 = \mu_1 \quad \text{vs} \quad H_1 : \mu_0 \neq \mu_1$$

Antes de efetuarmos o teste de comparação para as médias, temos que testar a igualdade de variâncias das variáveis entre os alunos do nível 0 ( $\sigma_0^2$ ) e os alunos do nível 1 ( $\sigma_1^2$ ), usando o teste F:

$$H_0 : \sigma_0^2 = \sigma_1^2 \quad \text{vs} \quad H_1 : \sigma_0^2 \neq \sigma_1^2$$

**Tabela 18 | Teste F e Teste t entre as variáveis resposta CN e MAT e as variáveis explicativas contínuas ao nível do aluno**

<i>Variáveis Explicativas Contínuas</i>	<i>CN</i>		<i>MAT</i>	
	<i>Teste F</i>	<i>Teste t</i>	<i>Teste F</i>	<i>Teste t</i>
	p – value	p – value	p - value	p – value
<i>AGE</i>	0.745	0.622	0.830	0.745
<i>ESTC</i>	0.008	0.000	0.012	0.000
<i>ESTE</i>	0.024	0.000	0.004	0.000
<i>ESTS</i>	0.000	0.000	0.000	0.000
<i>ESTCR</i>	0.000	0.000	0.000	0.000
<i>ESTM</i>	0.027	0.000	0.068	0.000
<i>ESCS</i>	0.000	0.000	0.000	0.000

Pela análise da Tabela 18, somente a variável *AGE* não se rejeita a hipótese nula da igualdade das médias a um nível de significância de 5%, logo há evidência estatística que as médias de idades entre os alunos do nível de desempenho 0 à disciplina de ciências e os alunos do nível de desempenho 1 à disciplina de ciências são iguais. Nas restantes variáveis não se rejeita a hipótese nula a um nível de significância de 5%.

No que diz respeito à variável dependente MAT, os resultados obtidos foram semelhantes: apenas para a variável AGE não se rejeita a hipótese nula da igualdade das médias a um nível de significância de 5%.

De seguida são apresentados alguns gráficos relativos às variáveis explicativas, para melhor compreender o nosso estudo.

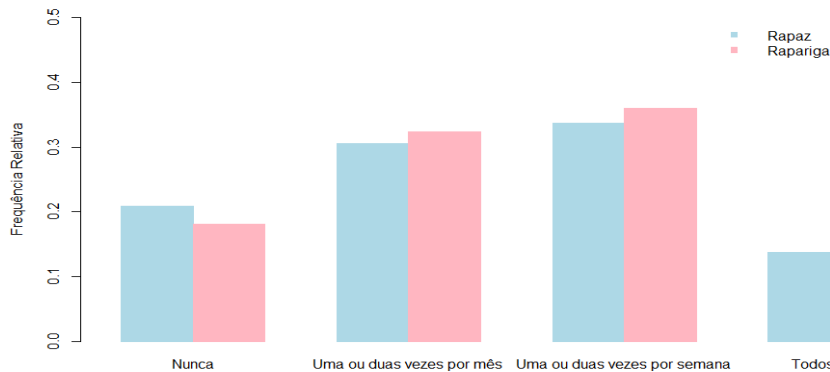


Figura 2 | Ajuda nos trabalhos de casa vs género do aluno

Na Figura 2 observa-se que um maior número de rapazes e raparigas têm ajuda nos trabalhos de casa pelos pais, uma ou duas vezes por semana, não se observando grandes diferenças entre géneros.

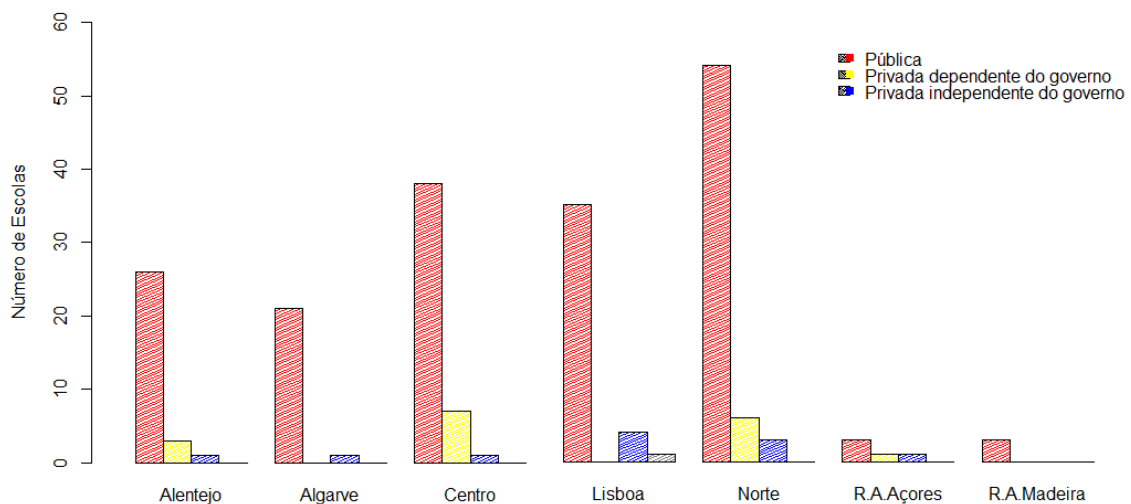
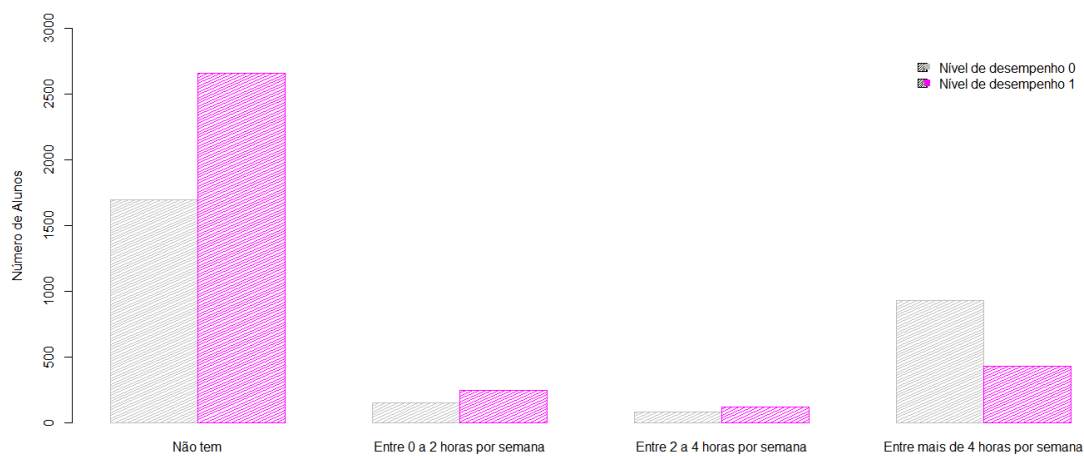


Figura 3 | Tipo de escola vs divisão regional

A Figura 3 sugere que é na região Norte que se situam o maior número de escolas públicas, seguida da região Centro. Relativamente às escolas privadas dependente do governo, estas também se encontram maioritariamente na região Centro e Norte. Quanto às escolas privadas independentes do governo, estas localizam-se em maior número na região Norte seguida da região autónoma dos Açores.

### 5.2.1 | Relação entre as variáveis resposta e as variáveis explicativas



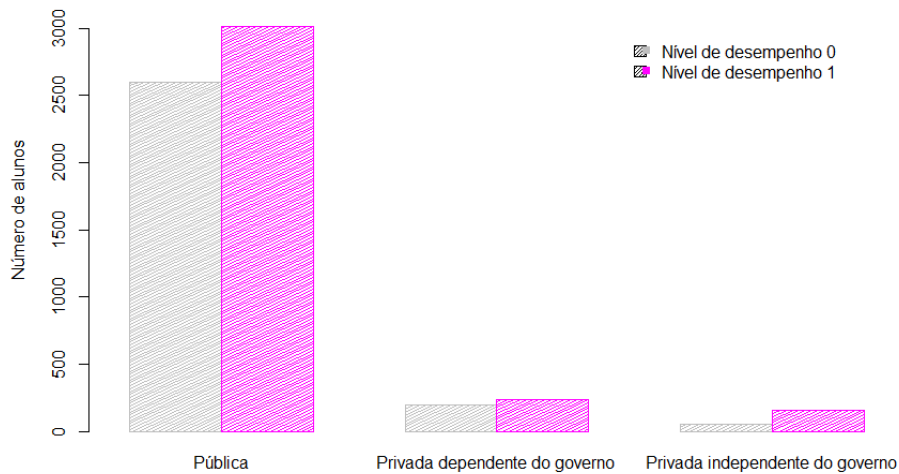
**Figura 4 | Níveis de desempenho a ciências vs horas de explicação a ciências**

A Figura 4, esta indica que a maioria dos alunos que se encontram no nível de desempenho 1 a ciências não têm explicações e os alunos que têm mais de 4 horas de explicação por semana são em maior número alunos com nível de desempenho 0 a ciências.



**Figura 5 | Níveis de desempenho a matemática vs horas de explicação a matemática**

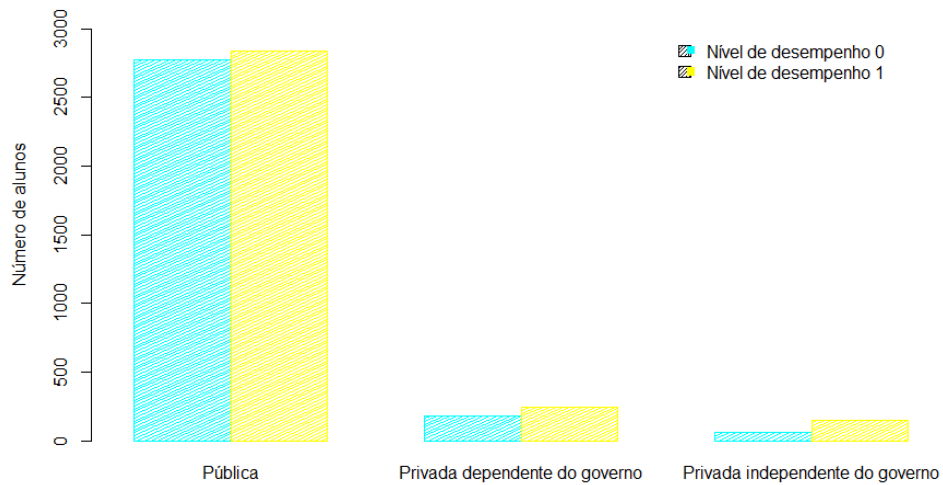
A Figura 5 mostra uma situação idêntica no caso da matemática. Isto é, a maioria dos alunos que se encontram no nível de desempenho 1 a matemática não têm explicações a matemática e os alunos que têm mais de 4 horas de explicação por semana a matemática são em maior número alunos com nível de desempenho 0.



**Figura 6 | Níveis de desempenho a ciências vs tipo de escola**

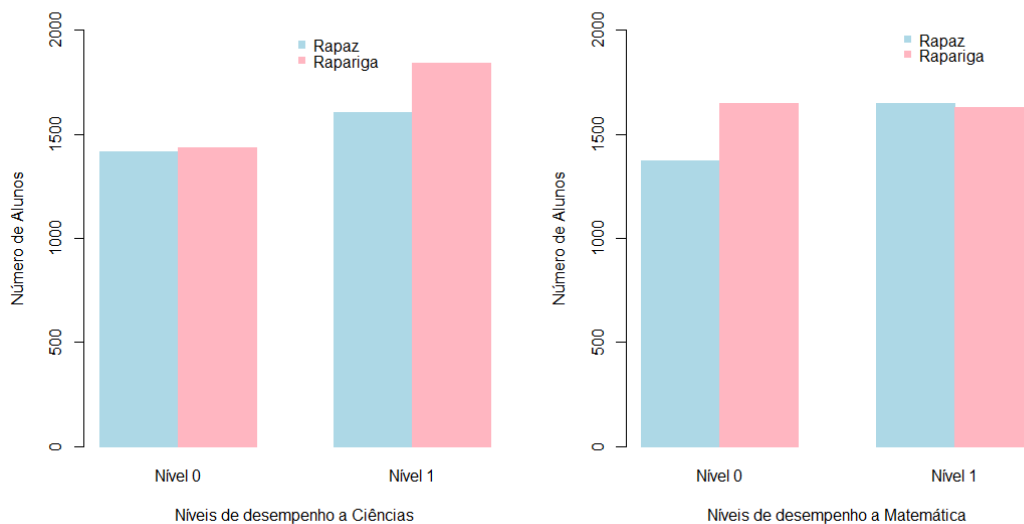
No nosso estudo, em todos os tipos de escolas encontram-se em maior número alunos no nível de desempenho 1 em ciências (ver Figura 6).





**Figura 7 | Níveis de desempenho a matemática vs tipo de escola**

No nosso estudo, em todos os tipos de escolas encontram-se em maior número alunos no nível de desempenho 1 em matemática (ver Figura 7).



**Figura 8 | Níveis de desempenho a ciências e matemática vs género do aluno**

A Figura 8 sugere que na área das ciências, as raparigas encontram-se maioritariamente no nível de desempenho 1. No entanto na área da matemática os rapazes encontram-se maioritariamente no nível de desempenho 1.



## Capítulo VI

### Apresentação dos Resultados

#### 6.1 | Apresentação dos resultados

Este capítulo divide-se em duas partes: a primeira parte refere-se à apresentação de resultados e respetivas conclusões relativas ao desempenho dos alunos portugueses a ciências e a segunda parte refere-se à apresentação de resultados e conclusões relativos ao desempenho dos alunos portugueses a matemática.

Neste trabalho foram utilizados os *softwares* estatísticos SPSS e R.

No *software* R pode-se ajustar um modelo de regressão multinível logística usando a função *lmer* que está disponível na *package lmer4*.

#### 6.2 | Resultados do desempenho dos alunos Portugueses a ciências para a variável resposta (CN)

##### 6.2.1 | Modelo Nulo

Em primeiro lugar estima-se o modelo apenas com o termo independente, modelo nulo definido em (1).

Tabela 19 | Estimativa da variância (modelo nulo)

<i>Variância do termo independente</i>	<i>Estimativa</i>	<i>Erro padrão</i>
$\left( \hat{\sigma}_{u0}^2 \right)$	1.520	0.723

A estimativa da variância entre escolas ( $\hat{\sigma}_{u0}^2$ ) no modelo nulo é de 1.520.

A partir dos resultados obtidos da Tabela 19, testa-se a hipótese nula  $H_0 : \sigma_{u0}^2 = 0$  efetuando um teste Wald. Como o *p-value* obtido é de 0.019, rejeita-se  $H_0$  ou seja há evidência estatística que a variância entre escolas é significativamente diferente de zero.

O coeficiente de correlação intra-classe (ICC) estimado neste modelo foi de 0.273, o que indica que 27.3% da variabilidade no desempenho dos alunos portugueses a ciências é explicada por diferenças entre as escolas que frequentam.

A Tabela 20 apresenta as estimativas do efeito fixo.

**Tabela 20| Estimativa do efeito fixo**

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio padrão</i>	<i>Valor-t</i>	<i>p-value</i>
$\left( \hat{\gamma}_{00} \right)$	0.117	0.095	1.232	0.196

No Anexo A2, os modelos de regressão logística simples multinível para cada uma das variáveis explicativas ao nível do aluno são apresentados.

### 6.2.2| Modelo de efeitos fixos com variáveis explicativas ao nível do aluno

Após a elaboração do modelo multinível simples da qual foram selecionadas as variáveis explicativas estatisticamente significativas procedeu-se à construção do modelo de efeitos fixos definido em (2).

Foi estimado o modelo de efeitos fixos com todas as variáveis ao nível do aluno e verificou-se que a variável *ESTE* (índice do uso de estratégias de elaboração) não se apresentava significativa, retirando-se esta variável do modelo.

De seguida elaborou-se o quadro com as estimativas, erros padrões, valores da estatística de teste, respetivos *p-value*, a razão de chance (OR) e os intervalos de confiança a 95% dos parâmetros do modelo de efeitos fixos com variáveis explicativas ao nível do aluno.

Tabela 21 | Modelo de efeitos fixos com variáveis explicativas ao nível do aluno

<i>Coefficientes</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Valor - t</i>	<i>p - value</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Intercept</b>	-1.145	0.204	-5.619	0.000 ***		
<b>IMI (ref: Imigrante)</b>						
Nativo	0.837	0.145	5.770	0.000 ***	2.309	[1.739,3.073]
<b>NEF (ref:3º Ciclo do Ensino Secundário)</b>						
Ensino secundário - curso geral/científico - humanístico	0.614	0.127	4.820	0.000***	1.848	[1.439,2.371]
Ensino secundário - curso tecnológico ou profissional/ formação profissional	-0.632	0.138	-4.584	0.000***	0.532	[0.406,0.696]
<b>TPC (ref: Nunca)</b>						
Uma a duas vezes por mês	-0.026	0.104	-0.246	0.806	0.974	[0.795,1.195]
Uma a duas vezes por semana	-0.121	0.103	-1.187	0.235	0.886	[0.724,1.083]
Todos os dias	-0.318	0.129	-2.461	0.014*	0.728	[0.565,0.937]
Desconhecido	-0.753	0.488	-1.544	0.123	0.471	[0.181,1.225]
<b>HEXPCN (ref: Nunca)</b>						
Tem entre 0 a 2 horas por semana	-0.169	0.157	-1.078	0.281	0.845	[0.622,1.149]
Tem entre 2 a 4 horas por semana	-0.407	0.204	-1.993	0.046*	0.666	[0.448,0.996]
Tem mais de 4 horas por semana	-0.989	0.090	-10.981	0.000***	0.372	[0.312,0.444]
<b>EXPAL (ref: Licenciatura ou equivalente)</b>						
3º Ciclo do Ensino Básico	-0.532	0.277	-1.920	0.055.	0.587	[0.339,1.006]
Ensino secundário - curso geral/científico - humanístico	-0.511	0.105	-4.879	0.000***	0.600	[0.490,0.738]
Ensino secundário - curso tecnológico ou profissional/ formação profissional	-0.331	0.092	-3.599	0.000***	0.718	[0.599,0.860]
<b>ESTC</b>	0.463	0.047	9.955	0.000***	1.589	[1.451,1.741]
<b>ESTS</b>	0.376	0.041	9.167	0.000***	1.457	[1.344,1.578]
<b>ESTM</b>	-0.400	0.043	-9.267	0.000 ***	0.670	[1.236,1.464]
<b>ESTCR</b>	0.297	0.039		0.000 ***	1.346	[1.246,1.452]
<b>REP (ref: Repetente)</b>						
Não Repetente	1.464	0.114	12.898	0.000 ***	4.323	[3.460,5.398]
<b>ESCS</b>	0.303	0.037	8.275	0.000 ***	1.354	[1.260,1.455]

**Desvio = 5279**

**ICC = 0.151**

**$\hat{\sigma}_{\mu 0}^2 = 0.339$**

Analisando a Tabela 21:

- \* Estima-se que a chance de um aluno nascido em Portugal pertencer à categoria de melhor desempenho é 2.309 vezes superior à chance de um aluno imigrante, quando os alunos são semelhantes nas restantes variáveis;

- \* Estima-se que a chance de um aluno frequentar o ensino secundário – curso geral/científico - humanístico, de pertencer à categoria de melhor desempenho é 1.848 vezes superior quando comparado com um aluno que frequente o 3º Ciclo do Ensino Básico, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno que frequente o ensino secundário – curso tecnológico ou profissional/ formação profissional, pertencer à categoria de melhor desempenho decresce cerca de 47% comparativamente a um aluno que frequente o 3º Ciclo do Ensino Básico, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que um aluno que faz os trabalhos de casa todos os dias, com a ajuda dos pais, pertencer à categoria de melhor desempenho decresce cerca de 27% comparativamente a um aluno que nunca faz os trabalhos de casa, com a ajuda dos pais, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que um aluno que tem mais de 4 horas de explicação por semana pertencer à categoria de melhor desempenho decresce cerca de 63% relativamente a um aluno que não tem explicações, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que um aluno com expectativas de terminar apenas o 3º Ciclo do Ensino Básico pertencer à categoria de melhor desempenho decresce cerca de 41% quando comparado com um aluno que pretenda tirar uma Licenciatura (ou equivalente), quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice de estratégias de controlo (*ESTC*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.589 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice de estratégias de sumarização (*ESTS*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.457 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice de estratégias de memorização (*ESTM*), estima-se que a chance de um aluno pertencer à categoria de melhor desempenho decresce cerca de 33%, quando os alunos são semelhantes nas restantes variáveis;

- \* Por cada aumento de uma unidade no índice de estratégias para compreender e lembrar (*ESTCR*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.346 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice económico, social e cultural (*ESCS*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.354 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* A estimativa da variância entre escolas para o modelo de efeitos fixos com variáveis explicativas ao nível do aluno é de 0.339.
- \* Como era de esperar o valor do ICC (correlação intra-classe) estimado neste modelo, reduz de 0.273 para 0.151, o que indica que quando são consideradas apenas as características do aluno, 15% da variabilidade no desempenho dos alunos portugueses a ciências é explicada por diferenças entre as escolas que frequentam;
- \* O coeficiente de determinação McFadden é  $R_{MF}^2 = 0.305$ , o coeficiente de determinação generalizado de Cox e Snell é  $R_{CS}^2 = 0.308$  e o coeficiente ajustado de Nagelkerke é  $R_N^2 = 0.440$ .

### **6.2.3| Modelo de efeitos fixos com variáveis explicativas ao nível do aluno e ao nível da Escola**

Após estimar o modelo de efeitos fixos com variáveis explicativas ao nível do aluno, procedeu-se à estimação do modelo de efeitos fixos com variáveis explicativas ao nível do aluno e ao nível da escola definido em (3).

As variáveis ao nível da escola foram introduzidas uma a uma e verificadas se eram estatisticamente significativas ( $p$ -value = 5%), quando estas não eram estatisticamente significativas eram retiradas do modelo. As variáveis estatisticamente significativas foram: *TIPO* (tipo de escola que o aluno frequenta), *ESCSMED* (média do índice económico, social e cultural), *COMPA* (comportamento do aluno), *COMP* (comportamento do professor).

De seguida elaborou-se o quadro com as estimativas, erros padrões, valores da estatística de teste, respetivos  $p$ -value, a razão de chance (OR) e os intervalos de confiança a 95% dos

parâmetros do modelo de efeitos fixos com variáveis explicativas ao nível do aluno e ao nível da escola (Tabela 22).

Tabela 22 | Modelo de efeitos fixos com variáveis explicativas ao nível do aluno e da escola

<i>Coefficientes</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Valor - t</i>	<i>p - value</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Intercept</b>	-0.790	0.179	-4.421	0.000 ***		
<b>IMI (ref: Imigrante)</b>						
Nativo	0.758	0.140	5.416	0.000 ***	2.134	[1.622,2.806]
<b>NEF (ref:3º Ciclo do Ensino Secundário)</b>						
Ensino secundário - curso geral/científico - humanístico	0.424	0.121	3.497	0.001 ***	1.528	[1.205,1.938]
Ensino secundário - curso tecnológico ou profissional/ formação profissional	-0.667	0.132	-5.070	0.000 ***	0.513	[0.397,0.664]
<b>TPC (ref: Nunca)</b>						
Uma a duas vezes por mês	0.018	0.101	0.176	0.861	1.018	[0.835,1.240]
Uma a duas vezes por semana	-0.081	0.099	-0.816	0.414	0.922	[0.759,1.120]
Todos os dias	-0.253	0.125	-2.034	0.042 *	0.777	[0.608,0.991]
Desconhecido	-0.701	0.473	-1.480	0.139	0.496	[0.196,1.255]
<b>HEXPCN (ref: Nunca)</b>						
Tem entre 0 a 2 horas por semana	-0.144	0.152	-0.945	0.344	0.866	[0.643,1.167]
Tem entre 2 a 4 horas por semana	-0.398	0.199	-2.004	0.045 *	0.672	[0.455,0.991]
Tem mais de 4 horas por semana	-0.970	0.087	-11.097	0.000 ***	0.379	[0.320,0.450]
<b>EXPAL (ref: Licenciatura ou equivalente)</b>						
3º Ciclo do Ensino Básico	-0.483	0.269	-1.795	0.073 .	0.617	[0.364,1.045]
Ensino secundário - curso geral/científico - humanístico	-0.463	0.101	-4.585	0.000 ***	0.629	[0.516,0.767]
Ensino secundário - curso tecnológico ou profissional/ formação profissional	-0.329	0.089	-3.702	0.000***	0.720	[0.604,0.856]
<b>ESTC</b>	0.444	0.045	9.832	0.000 ***	1.559	[1.427,1.703]
<b>ESTS</b>	0.354	0.040	8.921	0.000 ***	1.425	[1.319,1.541]
<b>ESTM</b>	-0.388	0.042	-9.268	0.000 ***	0.678	[0.625,0.737]
<b>ESTCR</b>	0.289	0.038	7.629	0.000 ***	1.335	[1.240,1.438]
<b>REP (ref: Repetente)</b>						
Não Repetente	1.317	0.109	12.090	0.000 ***	3.732	[3.014,4.620]
<b>ESCS</b>	0.269	0.037	7.316	0.000 ***	1.309	[1.218,1.407]
<b>TIPO (ref: Escola Pública)</b>						
Escola Privada dependente do governo	0.336	0.160	2.102	0.036 *	1.399	[1.023,1.916]
Escola privada independente do governo	0.196	0.275	0.711	0.477	1.217	[0.709,2.086]
<b>ESCSMED</b>	0.303	0.085	3.543	0.000 ***	1.354	[1.145,1.600]
<b>COMPA</b>	-0.129	0.052	-2.456	0.014 *	0.879	[0.794,0.974]
<b>COMP P</b>	0.129	0.052	2.458	0.014 *	1.138	[1.026,1.260]

**Desvio = 5288**  
**ICC = 0.084**  
 $\hat{\sigma}_{u0}^2 = 0.091$

Analisando a Tabela 22:



- \* Observa-se que em relação às variáveis explicativas ao nível do aluno todos os coeficientes estimados não se alteram muito quando introduzidas as variáveis explicativas ao nível da escola, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno que frequenta uma escola privada dependente do governo se encontrar na categoria de melhor desempenho seja 1.399 vezes superior quando comparado com a chance de um aluno que frequenta uma escola pública, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade na média do índice económico, social e cultural da escola (*ESCSMED*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.354 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no comportamento do aluno (*COMPA*), estima-se que a chance de um aluno pertencer à categoria de melhor desempenho decresce cerca de 12%, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no comportamento do professor (*COMPP*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.138 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* A estimativa da variância entre escolas para o modelo de efeitos fixos com variáveis explicativas ao nível do aluno e da escola é de 0.091.
- \* Como era de esperar o valor do ICC (correlação intra-classe) estimado neste modelo, reduz de 0.151 para 0.084, o que indica que 8% da variabilidade no desempenho dos alunos portugueses a ciências é explicada por diferenças entre as escolas que frequentam;
- \* O coeficiente de determinação McFadden é  $R_{MF}^2 = 0.304$ , o coeficiente de determinação generalizado de Cox e Snell é  $R_{CS}^2 = 0.317$  e o coeficiente ajustado de Nagelkerke é  $R_N^2 = 0.438$ .

### 6.3 | Resultados do desempenho dos alunos Portugueses a matemática para a variável resposta (MAT)

#### 6.3.1 | Modelo Nulo

Em primeiro lugar estima-se o modelo apenas com o termo independente, modelo nulo como definida em (1).

Tabela 23 | Estimativa da variância (modelo nulo)

<i>Variância do termo independente</i>	<i>Estimativa</i>	<i>Erro padrão</i>
$\left( \hat{\sigma}_{u0}^2 \right)$	1.680	0.844

A estimativa da variância entre escolas ( $\hat{\sigma}_{u0}^2$ ) no modelo nulo é de 1.680.

A partir dos resultados obtidos da Tabela 23, testa-se a hipótese nula  $H_0 : \sigma_{u0}^2 = 0$  efetuando um teste Wald. Como o *p-value* obtido é de 0.023, rejeita-se  $H_0$  ou seja há evidência estatística que a variância entre escolas é significativamente diferente de zero.

O coeficiente de correlação intra-classe (ICC) estimado neste modelo foi de 0.283, o que indica que 28.3% da variabilidade no desempenho dos alunos portugueses a matemática é explicada por diferenças entre as escolas que frequentam.

A Tabela 24 apresenta as estimativas do efeito fixo.

Tabela 24 | Estimativa do efeito fixo

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio padrão</i>	<i>Valor-t</i>	<i>p-value</i>
$\left( \hat{\gamma}_{00} \right)$	0.024	0.095	0.253	0.801

No Anexo A2 os modelos de regressão logística simples multinível para cada uma das variáveis explicativas ao nível do aluno são apresentados.

### **6.3.2| Modelo de efeitos fixos com variáveis explicativas ao nível do aluno**

Após a elaboração do modelo multinível simples da qual foram selecionadas as variáveis explicativas estatisticamente significativas procedeu-se à construção do modelo de efeitos fixos como definido em (2).

Com a elaboração do modelo de efeitos fixos com variáveis ao nível do aluno verificou-se que as variáveis *AGE* (idade do aluno) e *REG* (*região do aluno*) não se apresentavam significativa, retirando-se estas variáveis do modelo.

De seguida elaborou-se o quadro com as estimativas, erros padrões, valores da estatística de teste, respetivos *p-value*, a razão de chance (OR) e os intervalos de confiança a 95% dos parâmetros do modelo de efeitos fixos com variáveis explicativas ao nível do aluno (Tabela 25).

Tabela 25 | Modelo de efeitos fixos com variáveis explicativas ao nível do aluno

<i>Coefficientes</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Valor - t</i>	<i>p - value</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Intercept</b>	-1.469	0.203	-7.254	0.000 ***		
<b>SEX (ref: Feminino)</b>						
Masculino	1.232	0.083	14.809	0.000 ***	3.428	[2.912,4.035]
<b>IMI (ref: Imigrante)</b>						
Nativo	0.383	0.151	2.526	0.012 *	1.467	[1.090,1.974]
<b>NEF (ref:3º Ciclo do Ensino Secundário)</b>						
Ensino secundário- curso geral/científico - humanístico	0.742	0.132	5.578	0.000***	2.100	[1.618,2.724]
Ensino secundário - curso tecnológico ou profissional/ formação profissional	-0.315	0.145	-2.171	0.030 *	0.730	[0.549,0.970]
<b>TPC (ref: Nunca)</b>						
Uma a duas vezes por mês	-0.043	0.107	-0.403	0.687	0.958	[0.777,1.181]
Uma a duas vezes por semana	-0.159	0.105	-1.505	0.132	0.853	[0.694,1.049]
Todos os dias	-0.463	0.133	-3.478	0.001 ***	0.629	[0.485,0.817]
Desconhecido	-0.479	0.512	-0.937	0.349	0.619	[0.227,1.688]
<b>HEXPMAT (ref: Nunca)</b>						
Tem entre 0 a 2 horas por semana	-0.372	0.099	-3.756	0.000 ***	0.689	[0.568,0.837]
Tem entre 2 a 4 horas por semana	-0.613	0.121	-5.051	0.000 ***	0.542	[0.427,0.687]
Tem mais de 4 horas por semana	-1.239	0.106	-11.706	0.000 ***	0.290	[0.236,0.357]
<b>EXPAL (ref: Licenciatura ou equivalente)</b>						
<b>3º Ciclo do Ensino Básico</b>						
Ensino secundário - curso geral/científico - humanístico	-0.322	0.283	-1.137	0.256	0.725	[0.416,1.263]
Ensino secundário - curso tecnológico ou profissional/ formação profissional	-0.667	0.110	-6.072	0.000 ***	0.513	[0.414,0.637]
	-0.514	0.096	-5.331	0.000 ***	0.598	[0.495,0.723]
<b>ESTC</b>	0.281	0.051	4.851	0.000 ***	1.325	[1.182,1.484]
<b>ESTS</b>	0.399	0.043	9.194	0.000 ***	1.490	[1.369,1.623]
<b>ESTE</b>	0.195	0.061	3.192	0.001 **	1.215	[1.078,1.370]
<b>ESTM</b>	-0.334	0.044	-7.669	0.000 ***	0.716	[0.657,0.780]
<b>ESTCR</b>	0.329	0.040	8.142	0.000 ***	1.390	[1.284,1.504]
<b>REP (ref: Repetente)</b>						
Não Repetente	1.714	0.118	14.576	0.000 ***	5.551	[4.409,6.991]
<b>ESCS</b>	0.379	0.038	9.980	0.000 ***	1.461	[1.356,1.574]

<b>Desvio = 5124</b>
<b>ICC = 0.186</b>
$\hat{\sigma}_{u0}^2 = 0.560$

Analisando a Tabela 25:

- \* Estima-se que a chance de um aluno do sexo masculino pertencer à categoria de melhor desempenho é 3.428 vezes superior à chance de um aluno do sexo feminino, quando os alunos são semelhantes nas restantes variáveis, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno nascido em Portugal pertencer à categoria de melhor desempenho é 1.467 vezes superior à chance de um aluno imigrante, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno frequentar o ensino secundário – curso geral/científico - humanístico de pertencer à categoria de melhor desempenho é 1.848 vezes superior quando comparado com um aluno que frequente o 3º Ciclo do Ensino Básico, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno pertencer à categoria de melhor desempenho de um aluno frequentar o ensino secundário – curso tecnológico ou profissional/ formação profissional decresce cerca de 27% comparativamente a um aluno que frequente o 3º Ciclo do Ensino Básico, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno que faz os trabalhos de casa todos os dias, com a ajuda dos pais, pertencer à categoria de melhor desempenho decresce cerca de 37% comparativamente com um aluno que nunca faz os trabalhos de casa, com a ajuda dos pais, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno que tem mais de 4 horas de explicação por semana pertencer à categoria de melhor desempenho decresce cerca de 71% quando comparado com um aluno que não tem explicações, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance um aluno com expectativas de terminar apenas o 3º Ciclo do Ensino Básico se encontrar na categoria de melhor desempenho decresce cerca de 27% comparativamente a um aluno que pretenda tirar uma Licenciatura (ou equivalente), quando os alunos são semelhantes nas restantes variáveis;

- \* Por cada aumento de uma unidade no índice de estratégias de controlo (*ESTC*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.325 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice de estratégias de sumarização (*ESTS*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.490 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice de estratégias de estruturação (*ESTE*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.215 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice de estratégias de memorização (*ESTM*), estima-se que a chance de um aluno pertencer à categoria de melhor desempenho decresce cerca de 28%, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice de estratégias para compreender e relembrar (*ESTCR*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.390 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno não repetente pertencer à categoria de melhor desempenho é 5.551 vezes superior à de um aluno repetente, quando os alunos são semelhantes nas restantes variáveis;
- \* Por cada aumento de uma unidade no índice económico, social e cultural (*ESCS*), estima-se que a chance de um aluno se encontrar na categoria de melhor desempenho é 1.461 vezes superior, quando os alunos são semelhantes nas restantes variáveis;
- \* A estimativa da variância entre escolas para o modelo de efeitos fixos com variáveis explicativas ao nível do aluno é de 0.560;
- \* Como era de esperar o valor do ICC (correlação intra-classe) estimado neste modelo, reduz de 0.283 para 0.186, o que indica que quando são consideradas apenas as características do aluno, 19% da variabilidade no desempenho dos alunos portugueses a matemática é explicada por diferenças entre as escolas que frequentam;

- \* O coeficiente de determinação McFadden é  $R_{MF}^2 = 0.321$ , o coeficiente de determinação generalizado de Cox e Snell é  $R_{CS}^2 = 0.319$  e o coeficiente ajustado de Nagelkerke é  $R_N^2 = 0.457$ .

### **6.3.3| Modelo de efeitos fixos com variáveis explicativas ao nível do aluno e ao nível da escola**

Após estimar o modelo de efeitos fixos com variáveis ao nível do aluno, procedeu-se à estimação do modelo de efeitos fixos com variáveis ao nível do aluno e ao nível da escola como definido em (3).

As variáveis ao nível da escola foram introduzidas uma e uma e verificadas se eram estatisticamente significativas ( $p$ -value = 5%), quando estas não eram estatisticamente significativas eram retiradas do modelo. As variáveis estatisticamente significativas foram: *TIPO* (tipo de escola que o aluno frequenta), *DIM* (dimensão da escola), *ESCSMED* (média do índice económico, social e cultural).

De seguida elaborou-se o quadro com as estimativas, erros padrões, valores da estatística de teste, respetivos  $p$ -value, a razão de chance (OR) e os intervalos de confiança a 95% dos parâmetros do modelo de efeitos fixos com variáveis explicativas ao nível do aluno e ao nível da escola (Tabela 26).

Tabela 26 | Modelo de Efeitos Fixos com variáveis explicativas ao nível do aluno e da escola

<i>Coefficientes</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Valor - t</i>	<i>p - value</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Intercept</b>	-1.340	0.198	-6.780	0.000 ***		
<b>SEX (ref: Feminino)</b>						
Masculino	1.136	0.080	14.249	0.000 ***	3.114	[2.664,3.642]
<b>IMI (ref: Imigrante)</b>						
Nativo	0.389	0.144	2.707	0.007**	1.476	[1.113,1.955]
<b>NEF (ref:3º Ciclo do Ensino Secundário)</b>						
Ensino secundário - curso geral/científico - humanístico	0.519	0.124	4.206	0.000 ***	1.680	[1.319,2.141]
Ensino secundário - curso tecnológico ou profissional/ formação profissional	-0.419	0.135	-3.107	0.002 **	0.658	[0.505,0.857]
<b>TPC (ref: Nunca)</b>						
Uma a duas vezes por mês	-0.010	0.102	-0.097	0.923	0.990	[0.810,1.210]
Uma a duas vezes por semana	-0.165	0.101	-1.635	0.102	0.848	[0.696,1.033]
Todos os dias	-0.417	0.127	-3.282	0.001 **	0.659	[0.514,0.845]
Desconhecido	-0.437	0.483	-0.905	0.366	0.646	[0.251,1.665]
<b>HEXPMAT (ref: Nunca)</b>						
Tem entre 0 a 2 horas por semana	-0.325	0.095	-3.436	0.001 ***	0.723	[0.600,0.870]
Tem entre 2 a 4 horas por semana	-0.567	0.116	-4.870	0.000 ***	0.567	[0.452,0.713]
Tem mais de 4 horas por semana	-1.159	0.101	-11.443	0.000 ***	0.314	[0.257,0.383]
<b>EXPAL (ref: Licenciatura ou equivalente)</b>						
3º Ciclo do Ensino Básico	-0.196	0.273	-0.719	0.472	0.822	[0.482,1.403]
Ensino secundário - curso geral/científico - humanístico	-0.584	0.105	-5.574	0.000 ***	0.556	[0.454,0.685]
Ensino secundário - curso tecnológico ou profissional/ formação profissional	-0.425	0.092	-4.605	0.000 ***	0.654	[0.546,0.784]
<b>ESTC</b>	0.254	0.056	4.558	0.000 ***	1.289	[1.156,1.437]
<b>ESTS</b>	0.379	0.042	9.114	0.000 ***	1.461	[1.346,1.584]
<b>ESTE</b>	0.173	0.059	2.955	0.003 **	1.188	[1.060,1.334]
<b>ESTM</b>	-0.300	0.042	-7.178	0.000 ***	0.741	[0.683,0.804]
<b>ESTCR</b>	0.310	0.039	8.003	0.000 ***	1.363	[1.264,1.471]
<b>REP (ref: Repetente)</b>						
Não Repetente	1.530	0.111	13.845	0.000 ***	4.618	[3.717,5.732]
<b>ESCS</b>	0.336	0.038	8.942	0.000 ***	1.399	[1.300,1.506]
<b>TIPO (ref: Escola Pública)</b>						
Escola Privada dependente do governo	0.519	0.162	3.199	0.001 **	1.680	[1.223,2.310]
Escola privada independente do governo	0.088	0.272	0.324	0.746	1.092	[0.641,1.860]
<b>DIM (ref: Pequena)</b>						
Média	0.224	0.095	2.358	0.018 *	1.251	[1.039,1.507]
Grande	0.377	0.180	2.097	0.036 *	1.457	[1.025,2.073]
<b>ESCSMED</b>	0.258	0.091	2.832	0.005 **	1.294	[1.083,1.546]
<b>Desvio = 5142</b>						
<b>ICC = 0.084</b>						
<b><math>\hat{\sigma}_{u0}^2 = 0.091</math></b>						



Analisando a Tabela 26:

- \* Observa-se que em relação às variáveis explicativas ao nível do aluno todos os coeficientes estimados não se alteram muito quando introduzidas as variáveis explicativas ao nível da escola, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno que frequente numa escola privada dependente do governo se encontrar na categoria de melhor desempenho seja 1.680 superior à chance de um aluno que frequente uma escola pública, quando os alunos são semelhantes nas restantes variáveis;
- \* Estima-se que a chance de um aluno que frequente uma escola de grande dimensão se encontrar na categoria de melhor desempenho seja 1.457 vezes superior à chance de um aluno que frequente uma escola de pequena dimensão, quando os alunos são semelhantes nas restantes variáveis;
- \* A estimativa da variância entre escolas para o modelo de efeitos fixos com variáveis explicativas ao nível do aluno e da escola é de 0.091.
- \* Como era de esperar o valor do ICC (correlação intra-classe) estimado neste modelo, reduz de 0.186 para 0.084, o que indica que 8% da variabilidade no desempenho dos alunos portugueses a matemática é explicada por diferenças entre as escolas que frequentam.
- \* O coeficiente de determinação McFadden é  $R_{MF}^2 = 0.318$ , o coeficiente de determinação generalizado de Cox e Snell é  $R_{CS}^2 = 0.317$  e o coeficiente ajustado de Nagelkerke é  $R_N^2 = 0.454$ .



## Capítulo VII

### Conclusão

O objetivo primordial desta dissertação é identificar alguns dos fatores que possam determinar os resultados escolares dos alunos e perceber se as escolas, enquanto organizações, também podem contribuir para o sucesso dos seus alunos, bem como identificar em que medida as variáveis consideradas, por aluno e por escola, explicam o desempenho dos alunos às disciplinas de matemática e ciências. Para realizarmos o nosso estudo foram utilizados os dados de PISA de 2009.

No nosso estudo foi utilizada a regressão logística multinível para estimar dois modelos, um com variáveis só ao nível do aluno (nível 1) e outro com variáveis ao nível do aluno (nível 1) e ao nível da escola (nível 2).

Os resultados do PISA 2009 presentes nesta dissertação sugerem que a maioria dos alunos nas disciplinas de ciências e matemática se encontram em níveis de desempenho mais elevados.

Os alunos do sexo masculino são os que tiram melhores notas a matemática, contrariamente, à disciplina de ciências são os alunos do sexo feminino.

Na estimação do modelo com apenas variáveis ao nível do aluno concluiu-se que as variáveis explicativas que influenciam positivamente o desempenho dos alunos são muito semelhante em ambas as disciplinas: IMI (se o aluno nasceu em Portugal), ESTC (estratégias de controlo), ESTS (estratégias de sumarização), ESTCR (índice do uso de estratégias de compreender e relembrar), REP (se o aluno é repetente) e ESCS (índice económico, social e cultural). Há exceção do sexo do aluno que tem uma influência positiva somente à disciplina de matemática.

Em seguida, foi estimado o modelo de regressão logística multinível com variáveis explicativas ao nível do aluno e da escola, na medida em que este tipo de modelos é considerado adequado para determinar os impactos no desempenho escolar de alunos pertencentes a várias escolas.

No que diz respeito às variáveis explicativas ao nível do aluno, em ambas as disciplinas, estas não sofreram grandes alterações. Relativamente às variáveis ao nível da escola, as que vão influenciar o desempenho positivo à disciplina de ciências são o tipo de escola que os alunos frequentam, a média do índice económico, social e cultural e o comportamento do professor. Para a disciplina de ciências as que influenciam os alunos positivamente são o tipo de escola que os alunos frequentam, a dimensão da escola que os alunos frequentam e a média do índice económico, social e cultural.

Para ambas as disciplinas, o valor da variância entre escolas estimada e o valor do coeficiente intra-classe (ICC) vão diminuindo após a introdução de variáveis explicativas no modelo.

Para ambas as disciplinas, no modelo com variáveis explicativas ao nível do aluno e da escola, o coeficiente de correlação intra-classe (ICC) é de 0.084, o que indica que apenas 8% da variabilidade no desempenho escolar dos alunos é explicada por diferenças entre as escolas que os alunos frequentam.

Podemos concluir então que, em ambas as disciplinas, a chance dos alunos nativos se encontrarem na categoria de melhor desempenho é superior à chance dos alunos imigrantes e os alunos repetentes têm menor chance de se encontrar na categoria de melhor desempenho comparativamente aos alunos não repetentes.

Quanto à ajuda dos pais nos trabalhos de casa podemos concluir que esta tem um impacto negativo no desempenho dos alunos, a sua razão de chance é maior quanto menor a ajuda dos pais no trabalho de casa.

Relativamente à variável número de horas de explicação, os modelos sugerem que o tempo que os alunos levam a estudar nas aulas de apoio/ explicações após a escola tem uma relação negativa com o bom desempenho (como foi analisado no capítulo dos resultados). Isto acontece porque os estudantes que frequentam essas aulas fazem isso apenas como recuperação e não e para aprimorar os estudos.

Quanto às expectativas dos alunos, na sua maioria os alunos que se encontram no nível de desempenho mais elevado esperam alcançar uma licenciatura (ou equivalente).

As estratégias de controlo, de estruturação, de compreender e relembrar têm um impacto positivo no desempenho escolar nos alunos portugueses, em ambas as disciplinas.

As estratégias de memorização tem um impacto negativo no desempenho escolar dos alunos portugueses, em ambas as disciplinas.

Um aluno português que frequente uma escola privada dependente do governo tem maior chance de se encontrar na categoria de melhor desempenho comparativamente a um aluno português que frequente uma escola pública.

A média do índice económico, social e cultural influencia positivamente o desempenho dos alunos portugueses em ambas as disciplinas.

Na disciplina de ciência o comportamento do aluno tem um impacto negativo e o comportamento do professor tem um impacto positivo no desempenho do aluno português.

Na disciplina de matemática, um aluno que frequenta uma escola de grandes dimensões tem maior chance de se encontrar na categoria de melhor desempenho quando comparado com um aluno que frequente uma escola de pequenas dimensões.

## **7.1 | Sugestões para trabalhos futuros**

Penso ser pertinente para trabalhos futuros comparar estas conclusões com os dados do PISA de 2012 cujos resultados e bases de dados foram divulgados em 2013 mas já tínhamos iniciado o estudo dos dados do PISA 2009 há já algum tempo.

O estudo de modelos de regressão logística multinível de efeitos aleatórios será um tema interessante a desenvolver.

Também poderá ser interessante a introdução de outras variáveis explicativas que não foram consideradas neste trabalho para verificar até que ponto a introdução dessas novas variáveis poderá ou não influenciar o desempenho dos alunos.



## Bibliografia

- \* Martins, A., Cabrita, I. (1991). A Problemática do Insucesso Escolar, Universidade de Aveiro.
- \* Albernaz, A., Ferreira, F., Franco, C. (2002). Qualidade e equidade no ensino fundamental brasileiro. *Pesquisa e Planejamento econômico*, Rio de Janeiro, (32):3.
- \* Aitkin, M., Anderson, D., Hinde J. (1981), Statistical modelling in school effectiveness studies (with discussion), *Journal of the Royal Statistical Society A*, 149, 1-43.
- \* Barbosa, M. E. , Fernandes, C. (2001). A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4ª série. Em C. Franco (Org.), *Avaliação, ciclos e promoção na educação*. Porto Alegre: Artmed.
- \* Barbosa, M. E. , Fernandes, C. (2000). Modelo multinível: Uma aplicação a dados de avaliação educacional. *Estudos em Avaliação Educacional*, 22, 135-153.
- \* Bennet, N. (1976), *Teaching Styles and Pupil Progress*, London: Open Books.
- \* Bickel, R. (2007). *Multilevel Analysis for Applied Research: It's Just Regression*. Guilford Press.
- \* Bratti, M., Checchi, D., Filippin, A. (2007). Territorial Differences in Italian Students' Mathematical Competencies: Evidence from PISA 2003, IZA Discussion Papers 2603, Institute for the Study of Labor (IZA).
- \* Coleman, S., Campbell, E., Hobson, C., McPartland, F., Mood, A., Weinfeld, F., York, R. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- \* Callens, M., Croux, C. (2005). Performance of likelihood-based estimation methods for multilevel binary regression models. *Journal of Statistical Computation and Simulation*. 75 (12), 1003–1017.
- \* Carneiro, P. (2008). Equality of opportunity and educational achievement in Portugal, *Portuguese Economic Journal*, 7(1), 17-41.
- \* Carneiro, P., Reis, H. (2009). Sources of inequality in educational achievement: an international comparison, UCL manuscript.

- \* Catalán-Reyes, M., Galindo-Villardón, M. (2003). Utilización de modelo multinivel en investigación sanitaria. *Gaceta Sanitaria*, 17 (3): 35-52.
- \* Collet, D. (1952). *Modelling Binary Data*. Texts in Statistical Science. 2nd Chapman & Hall/CRC.
- \* Dias, V., Ferrão, M. (2006). Modelo Multinível do Desempenho Escolar de Alunos Socialmente Desfavorecidos em Escolas Públicas/Privadas – Aplicação aos Dados Portugueses do PISA 2000. *Psicologia e Educação* 2.
- \* Diggle, P. (2002). *Analysis of Longitudinal Data*. 2nd Oxford.
- \* Diogo, A. (1998). *Famílias Escolaridade, representações parentais da escolaridade, classe social e dinâmica familiar*. Lisboa: Edições Colibri.
- \* Dronkers, J., Robert, P. (2003). *The effectiveness of public and private schools from a comparative perspective*. San Domenico di Fiesole (Fi): European University Institute. EUI working paper SPS 2003/13.
- \* Ferrão, M. (2003). *Introdução aos modelos de regressão multinível em educação*. Campinas: Editora Komedi.
- \* Ferrão, M., Dias, V. (2006). Modelos multinível do desempenho escolar dos alunos socialmente desfavorecidos em escolas público/privadas - aplicação aos dados portugueses de pisa 2000. *Revista brasileira de Economia*, 5(2), 63–77.
- \* Ferreira, J. (2004). *Planos Hierárquicos*. Trabalho no âmbito da disciplina de Seminário de Tópicos Complementares de Matemática do MEC. Lisboa, Universidade Aberta.
- \* Fukuda, C. (2003). *O ensino eficaz na educação básica: um modelo descritivo dos fatores de eficácia*. Tese de Doutorado, Universidade de Brasília, Brasília.
- \* Garcia, C. (1995). *Formação de Professores – para uma formação educativa*. Coleção Ciências da Educação, 2. Porto: Porto Editora.
- \* Garner, C., Raudenbush, S. (1991). Neighborhood Effects on Educational Attainment: A Multilevel Analysis. *Sociology of Education*, 64, 251-262.
- \* GAVE (2001). *PISA 2000 – Resultados do estudo internacional*, Lisboa: GAVE.
- \* GAVE (2002). *PISA 2000 – Conceitos fundamentais em jogo na avaliação da literacia científica e competências dos alunos portugueses*, Lisboa: GAVE.



## Bibliografia

- \* GAVE (2003). PISA 2000 – Conceitos fundamentais em jogo na avaliação da literacia científica e competências dos alunos portugueses, Lisboa: GAVE.
- \* GAVE (2004). PISA 2003 – Resultados do estudo internacional, Lisboa: GAVE.
- \* GAVE (2007). PISA 2006 – Competências científicas dos alunos portugueses, Lisboa: GAVE.
- \* GAVE (2010). PISA 2009 – Competências dos alunos portugueses: síntese de resultados, Lisboa: GAVE.
- \* Goldstein, H. (1995). Multilevel statistical models. 2nd Ed. London: Edward Arnold.
- \* Goldstein, H., Rasbash, J. (1996). Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society A*, 13, 159-505.
- \* Guo, G., Zhao, H. (2000), Multilevel modeling for binary data, *Annual Review of Sociology*, 26, 441-462.
- \* Gutierrez, G. (2005), Estimação das escalas dos construtos capital social, capital cultural e capital económico e análise do efeito escola nos dados do Peru-PISA 2000. Pontifícia Universidade Católica do Rio de Janeiro.
- \* Hesketh, S. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal*, 2(1), 1–21.
- \* Hox, J. (2002). Multilevel Analysis. Techniques and Application. Lawrence Erlbaum Associates.
- \* Leyland, A., Goldstein, H. (2001). Multilevel Modelling of Health Statistics. John Wiley & Sons.
- \* Murillo, F. (2003). Una panorámica de la investigación iberoamericana sobre eficacia escolar. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y cambio em Educación*, 1 (1), 1–14.
- \* OCDE (1999). Measuring Student Knowledge and Skills – A New Framework for Assessment, Paris: OCDE Publishing.
- \* OCDE (2003b). The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills, Paris: OCDE Publishing.
- \* OCDE (2004b). Assessment Scientific, Readins and Mathmatical Literacy – A Framework for PISA 2006, Paris: OCDE Publishing.

- \* OCDE (2006). *Assessing Scientific, Reading and Mathematical Literacy – A Framework for PISA 2006*, Paris: OCDE Publishing.
- \* OCDE (2007). *PISA 2006 Science Competencies for Tomorrow's World (Vol. I)*, Paris: OCDE Publishing.
- \* OCDE (2009a). *PISA 2006 Technical Report*, Paris: OCDE Publishing.
- \* OCDE (2009b). *Equally prepared for life? How 15-year-old boys and girls perform in school*, Paris: OCDE Publishing.
- \* OCDE (2009c). *PISA 2009 Assessment Framework – Key Competencies in Reading, Mathematics and Science*, Paris: OCDE Publishing.
- \* OCDE (2010a). *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (Volume I)*, OECD Publishing.
- \* OCDE (2010b). *PISA 2009 Results: Overcoming Social Background: Equity in Learning Opportunities and Outcomes (Volume II)*, OECD Publishing.
- \* OCDE (2010c). *PISA 2009 Results: Learning Trends: Changes in Student Performance Since 2000 (Volume V)*, PISA, OECD Publishing.
- \* OCDE (2010d). *PISA 2009 at a Glance*, Paris: OCDE Publishing.
- \* OCDE (2010e). Chapter 15. Proficiency Scale Construction in PISA 2009 Technical Report (preliminary version).
- \* OECD (2010). *PISA 2009 Results: What Students Know and Can Do, (1)*, Paris: Organization for Economic Co-operation and Development.
- \* Pereira, M., Reis, H., *Diferenças Regionais no Desempenho dos Alunos Portugueses: Evidência do programa PISA da OCDE*; Publicação.
- \* Pereira, M., (2011). Uma análise da evolução do desempenho dos estudantes portugueses no programme for international student assessment (pisa) da ocde. Banco de Portugal - Boletim económico, (II):131–146.
- \* Pierre, J., Huguet, D., Christine, B. (1994). Educação familiar e parental. *Revista Inovação*, 7, 289-305.
- \* Piscarreta, S. (2002). *Malmequer, bem-me-quer, muito pouco ou nada: Representações sociais da Matemática em alunos do 9º ano de escolaridade*. Dissertação de Mestrado, Lisboa, Universidade Aberta.

## Bibliografia

- \* Raudenbush, S., Bryk, A. (2002). Hierarchical linear models. Sage Publications, 2nd edition.
- \* Rodriguez, G., Goldman, N. (1995). An assessment of estimation procedures for multilevel models with binary response. *Journal of the Royal Statistical Society A*, 158, 73 – 89.
- \* Rodrigues, M. (2002). Instrumentos de avaliação educacional: uma visão pedagógica e psicométricas integradas estudo das provas do SAEB de Matemática 8ª série – 1997 e 1999. Dissertação de Mestrado, Universidade de Brasília, Brasília.
- \* Seabra, T. (1994), Estratégias familiares de socialização das crianças: etnicidade e classes sociais, Tese de Mestrado em Sociologia, Lisboa, Instituto Superior de Ciências do Trabalho e da Empresa.
- \* Skrandal A., Rabe-Hesketh, S. (2004). Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models. Chapman & Hall/CRC.
- \* Snijders, T., Bosker, R. (1999). Multilevel analysis: An introduction to basic and advanced multilevel modeling. London: SAGE Publications.
- \* Soares, J. F., César, C. C. & Mambrini, J. (2001). Determinantes de desempenho dos alunos do ensino básico brasileiro: evidências do SAEB de 1997. Em C. Franco (Org.), *Avaliação, ciclos e promoção na educação* (pp. 121-153). Porto Alegre: Artmed.
- \* Twisk, J. (2003). Applied Longitudinal Data Analysis for Epidemiology. Cambridge University Press.
- \* Twisk J. (2005). Applied Multilevel Analysis: a Practical Guide. Cambridge University Press.
- \* Venables, W., Ripley, B. (2002) Modern Applied Statistics. Statistics and Computing with S. 4nd Springer.
- \* Wössman, L. (2007). Fundamental Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries, IZA Discussion Papers 2880, Institute for the Study of Labor (IZA).



## Anexos

### A1 – CN - Regressão logística multinível simples

Tabela 27 | Regressão logística simples com a variável explicativa *SEX*

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>SEX (ref: Feminino)</i>	0.115	0.094	1.220	0.222		
Masculino	0.004	0.058	0.061	0.951	1.004	[0.896,1.124]

Tabela 28 | Regressão logística simples com a variável explicativa *AGE*

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>AGE</i>	0.825	0.978	0.843	0.399		
	-0.046	0.064	-0.727	0.467	0.955	[0.843,1.082]

Tabela 29 | Regressão logística simples com a variável explicativa *IMI*

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>IMI (ref: Imigrante)</i>	-0.621	0.143	-4.336	0.000 ***		
Nativo	0.795	0.120	6.642	0.000 ***	2.213	[1.751,2.798]

Tabela 30 | Regressão logística simples com a variável explicativa *NEF*

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>NEF (ref: 3º Ciclo do Ensino Básico)</i>	-0.869	0.077	-11.322	0.000 ***		
Ensino Secundário – curso geral/científico – humanístico	2.370	0.086	27.615	0.000 ***	10.697	[9.044,12.662]
Ensino Secundário – curso tecnológico ou profissional/formação profissional	-0.210	0.105	-2.011	0.044 *	0.811	[0.659,0.994]

Tabela 31 | Regressão logística simples com a variável explicativa *AGREG*

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>AGREG (ref: Família tradicional)</i>	0.175	0.090	1.944	0.052.		
Um parente	-0.039	0.078	-0.494	0.621	0.962	[0.825,1.122]
Outra estrutura	-1.151	0.195	-5.892	0.000 ***	0.316	[0.216,0.464]
Desconhecido	-1.928	0.404	-4.770	0.000 ***	0.146	[0.066,0.321]

Tabela 32 | Regressão logística simples com a variável explicativa TPC

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>TPC (ref: Nunca)</i>	0.105	0.107	0.984	0.325		
Uma a duas vezes por mês	0.166	0.084	1.977	0.048 *	1.181	[1.001,1.391]
Uma a duas vezes por semana	0.048	0.083	0.578	0.563	1.049	[0.892,1.233]
Todos os dias	-0.322	0.103	-3.127	0.002 ***	0.725	[0.591,0.886]
Desconhecido	-1.400	0.388	-3.606	0.000 ***	0.247	[0.115,0.527]

Tabela 33 | Regressão logística simples com a variável explicativa HEXPCN

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>HEXPCN (ref: Nunca)</i>	0.391	0.091	4.290	0.000 ***		
Tem entre 0 a 2 horas por semana	-0.078	0.123	-0.631	0.528	0.925	[0.726,1.177]
Tem entre 2 a 4 horas por semana	-0.265	0.171	-1.552	0.121	0.767	[0.549,1.072]
Tem mais de 4 horas por semana	-1.200	0.075	-16.106	0.000 ***	0.301	[0.260,0.348]

Tabela 34 | Regressão logística simples com a variável explicativa EXPAL

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>EXPAL (ref: Licenciatura ou equivalente)</i>	0.784	0.086	9.156	0.000 ***		
<i>3º Ciclo do Ensino Básico</i>	-1.508	0.229	-6.600	0.000 ***	0.221	[0.141,0.346]
Ensino secundário – curso geral/científico - humanístico	-1.003	0.087	-11.579	0.000 ***	0.367	[0.309,0.434]
Ensino secundário – curso tecnológico ou profissional/ formação profissional	-1.428	0.071	-19.997	0.000 ***	0.240	[0.208,0.275]

Tabela 35 | Regressão logística simples com a variável explicativa ESTC

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>ESTC</i>	0.060	0.085	0.708	0.479		
	0.666	0.035	18.794	0.000***	1.946	[0.735,1.815]

Tabela 36 | Regressão logística simples com a variável explicativa ESTS

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <i>ESTS</i>	0.138	0.080	1.732	0.083		
	0.834	0.034	24.699	0.000***	2.303	[2.155,2.460]

Tabela 37 | Regressão logística simples com a variável explicativa ESTE

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	-0.038	0.090	-0.421	0.674		
ESTE	0.419	0.036	11.526	0.000***	1.520	[1.415,1.632]

Tabela 38 | Regressão logística simples com a variável explicativa ESTM

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	0.072	0.091	0.787	0.431		
ESTM	-0.167	0.030	-5.526	0.000***	0.846	[0.798,0.898]

Tabela 39 | Regressão logística simples com a variável explicativa ESTCR

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	0.211	0.082	2.571	0.010*		
ESTCR	0.717	0.031	22.884	0.000***	2.049	[1.927,2.179]

Tabela 40 | Regressão logística simples com a variável explicativa REG

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	0.509	0.2015	2.529	0.012*		
REG (ref: Norte)						
Alentejo	-0.553	0.309	-1.792	0.073.	0.575	[0.314,1.053]
Algarve	-0.904	0.340	-2.657	0.008 **	0.405	[0.208,0.789]
Centro	-0.360	0.275	-1.308	0.191	0.698	[0.407,1.197]
Norte	-0.351	0.257	-1.370	0.171	0.704	[0.426,1.164]
R.A.Açores	-2.003	0.670	-2.989	0.003 **	0.135	[0.036,0.502]
R.A.Madeira	-0.876	0.742	-1.181	0.237	0.417	[0.097,1.782]
Desconhecido	0.455	0.591	0.771	0.441	1.576	[0.495,5.017]

Tabela 41 | Regressão logística simples com a variável explicativa REP

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	-1.270	0.077	-16.440	0.000 ***		
REP (ref: Imigrante)						
Nativo	2.420	0.078	31.020	0.000 ***	11.250	[9.654,13.108]

Tabela 42 | Regressão logística simples com a variável explicativa ESCS

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	0.333	0.079	4.216	0.000 ***		
ESCS	0.546	0.030	18.034	0.000 ***	1.726	[1.627,1.832]

## A2 – MAT - Regressão logística multinível simples

Tabela 43| Regressão logística simples com a variável explicativa *SEX*

<i>Variáveis</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Estatística de Teste</i>	<i>Valor p</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Constante</b> <i>SEX (ref: Feminino)</i>	-0.239	0.101	-2.365	0.018 *		
Masculino	0.445	0.059	7.570	0.000 ***	1.561	[1.391,1.751]

Tabela 44| Regressão logística simples com a variável explicativa *AGE*

<i>Variáveis</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Estatística de Teste</i>	<i>Valor p</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Constante</b> <i>AGE</i>	0.911	0.985	0.926	0.354		
	-0.061	0.064	-0.955	0.340	0.941	[0.830,1.067]

Tabela 45| Regressão logística simples com a variável explicativa *IMI*

<i>Variáveis</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Estatística de Teste</i>	<i>Valor p</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Constante</b> <i>IMI (ref: Imigrante)</i>	-0.502	0.146	-3.431	0.000 ***		
Nativo	0.515	0.120	4.287	0.000 ***	1.674	[1.322,2.116]

Tabela 46| Regressão logística simples com a variável explicativa *NEF*

<i>Variáveis</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Estatística de Teste</i>	<i>Valor p</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Constante</b> <i>NEF (ref: 3º Ciclo do Ensino Básico)</i>	-1.077	0.081	-13.287	0.000 ***		
Ensino Secundário – curso geral/científico – humanístico	2.344	0.087	26.939	0.000 ***	10.422	[8.784,12.354]
Ensino Secundário – curso tecnológico ou profissional/ formação profissional	0.139	0.106	1.317	0.188	1.149	[0.934, 1.414]

Tabela 47| Regressão logística simples com a variável explicativa *AGREG*

<i>Variáveis</i>	<i>Estimativa dos parâmetros</i>	<i>Desvio – Padrão</i>	<i>Estatística de Teste</i>	<i>Valor p</i>	<i>Razão de Chance (OR)</i>	<i>IC [95%] da Razão de Chance</i>
<b>Constante</b> <i>AGREG (ref: Família tradicional)</i>	0.065	0.094	0.685	0.493		
Um parente	-0.185	0.079	-2.341	0.019 *	0.831	[0.712,0.970]
Outra estrutura	-1.346	0.210	-6.406	0.000 ***	0.260	[0.173,0.393]
Desconhecido	-2.347	0.495	-4.745	0.000 ***	0.096	[0.036,0.252]



Tabela 48 | Regressão logística simples com a variável explicativa TPC

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <b>TPC (ref: Nunca)</b>	-0.002	0.111	-0.020	0.985		
Uma a duas vezes por mês	0.139	0.084	1.655	0.098	1.149	[0.975,1.356]
Uma a duas vezes por semana	-0.005	0.083	-0.058	0.954	0.995	[0.846,1.171]
Todos os dias	-0.387	0.105	-3.700	0.000 ***	0.679	[0.553,0.833]
Desconhecido	-1.239	0.388	-3.193	0.001 ***	0.289	[0.135,0.620]

Tabela 49 | Regressão logística simples com a variável explicativa HEXPMAT

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <b>HEXPMAT (ref: Nunca)</b>	0.304	0.098	3.097	0.002 **		
Tem entre 0 a 2 horas por semana	-0.316	0.079	-4.001	0.000 ***	0.729	[0.624,0.851]
Tem entre 2 a 4 horas por semana	-0.225	0.097	-2.320	0.020 *	0.799	[0.660,0.966]
Tem mais de 4 horas por semana	-1.354	0.086	-15.741	0.000 ***	0.258	[0.218,0.306]

Tabela 50 | Regressão logística simples com a variável explicativa EXPAL

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <b>EXPAL (ref: Licenciatura ou equivalente)</b>	0.597	0.091	6.582	0.000 ***		
3º Ciclo do Ensino Básico	-1.219	0.226	-5.394	0.000 ***	0.295	[0.190,0.460]
Ensino Secundário – curso geral/científico – humanístico	-0.993	0.088	-11.325	0.000 ***	0.371	[0.312,0.440]
Ensino Secundário – curso tecnológico ou profissional/ formação profissional	-1.340	0.072	-18.558	0.000 ***	0.262	[0.227,0.302]

Tabela 51 | Regressão logística simples com a variável explicativa ESTC

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <b>ESTC</b>	-0.075	0.089	-0.838	0.402		
	0.523	0.034	15.497	0.000 ***	1.687	[1.580,1.803]

Tabela 52 | Regressão logística simples com a variável explicativa ESTS

Variáveis	Estimativa dos parâmetros	Desvio - Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
<b>Constante</b> <b>ESTS</b>	-0.014	0.084	-0.160	0.873		
	0.722	0.033	21.800	0.000 ***	2.059	[1.929,2.197]

Tabela 53 | Regressão logística simples com a variável explicativa ESTE

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	
Constante	-0.185	0.095	-1.951	0.051 .		
ESTE	0.426	0.037	11.649	0.000 ***	1.531	[1.426,1.646]

Tabela 54 | Regressão logística simples com a variável explicativa ESTM

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	-0.076	0.096	-0.792	0.428		
ESTM	-0.188	0.030	-6.168	0.000 ***	0.829	[0.781,0.880]

Tabela 55 | Regressão logística simples com a variável explicativa ESTCR

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	0.052	0.086	0.605	0.545		
ESTCR	0.668	0.031	21.399	0.000 ***	1.950	[1.835,2.074]

Tabela 56 | Regressão logística simples com a variável explicativa REG

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	0.058	0.167	0.347	0.729		
REG (ref: Norte)						
Alentejo	-0.215	0.296	-0.727	0.467	0.807	[0.452,1.440]
Algarve	-0.760	0.332	-2.288	0.022 *	0.468	[0.244,0.897]
Centro	0.091	0.257	0.353	0.724	1.095	[0.662,1.811]
Lisboa	0.148	0.269	1.574	0.116	1.160	[0.685,1.964]
R.A.Açores	-1.513	0.677	0.552	0.581	0.220	[0.059,0.830]
R.A.Madeira	-0.962	0.779	-2.236	0.025 *	0.382	[0.083,1.758]
Desconhecido	0.951	0.604	-1.235	0.217	2.588	[0.792,8.456]

Tabela 57 | Regressão logística simples com a variável explicativa REP

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	-1.525	0.083	-18.380	0.000 ***		
REP						
Nativo	2.565	0.082	31.310	0.000 ***	13.000	[11.077,15.272]

Tabela 58 | Regressão logística simples com a variável explicativa ESCS

Variáveis	Estimativa dos parâmetros	Desvio – Padrão	Estatística de Teste	Valor p	Razão de Chance (OR)	IC [95%] da Razão de Chance
Constante	0.212	0.083	2.551	0.000 ***		
ESCS	0.612	0.031	19.765	0.000 ***	1.844	[1.736,1.960]